

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: May 21, 2004, 20:33:12 ; Search time 530 Seconds  
(without alignments)  
2605.027 Million cell updates/sec

Title: US-09-931-457A-31

Perfect score: 163  
Sequence: 1 MAVERSGIADVTLEIKTP.....LSVIFRSVREASMTFEP 325

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPTO\_spool/US09931457/rmat.18052004.121728.20289/app\_query.fasta.1.519  
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LIST=45 -DOCALLIGN=200 -STAR=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457 @CGN 1.1 470 @rmat.18052004.121728.20289 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOCG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THRAUS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1623	100.0	1362	2	AAV99906 Fragment
2	1431	88.2	1303	2	AAQ44450 Cysteine
3	1388	85.5	946	9	ADG76317 DNA homol
4	1384	85.3	1120	3	AAQ42360 Arabidops
5	1374	84.7	966	6	ABZ13088 Arabidops
6	1374	84.7	966	6	ADA68423 Arabidops
7	1360	83.8	993	7	ADA69412 Rice gene
8	1359	83.7	5259	7	ADA70437 Rice gene

9	1271	78.3	1483	2	AAQ74413 Spinach c
10	1261	77.7	1503	2	AAC34662 Arabidops
11	1255	77.3	1329	6	ABZ12294 Arabidops
12	1196	73.7	1224	3	AAC40774 Arabidops
13	1192	73.4	922	6	ABZ12690 Arabidops
14	1192	73.4	999	3	AAC45719 Arabidops
15	1192	73.4	1423	3	AAC41206 Arabidops
16	1191	73.0	1302	3	AAC7965 Arabidops
17	1184	73.0	1302	3	AAC33559 Arabidops
18	1179	72.6	975	9	ADG64221 Thalecres
19	1179	72.6	1200	3	AAC44727 Arabidops
20	1166	72.6	1302	3	AAC45323 Arabidops
21	1149	70.8	1215	3	AAC44993 Arabidops
22	1129	69.6	1037	9	ADG76318 DNA homol
23	1125	69.3	1006	9	ADG75697 DNA homol
24	1092	67.3	1030	9	ADG17590 DNA (Seq)
25	1077	66.4	725	9	ADD17591 DNA (Seq)
26	1065.5	65.7	918	3	AAC41635 Arabidops
27	1045	64.4	1481	3	AAC47004 Arabidops
28	1045	64.4	1529	4	ABA01246 O-acetyl
29	1017	62.7	1431	3	AAC46837 Arabidops
30	1006.5	62.0	1417	3	AAC51477 Arabidops
31	982	60.5	1089	3	AAC51458 Arabidops
32	980.5	60.4	683	9	ADG75696 DNA homol
33	974.5	60.0	1200	7	ADA69647 Rice gene
34	973	60.0	657	8	ACL17131 DNA clone
35	968.5	59.7	945	7	AAC24366 Prokaryot
36	962	59.3	913	9	ADG76310 DNA homol
37	962	58.3	933	7	ADD17569 DNA (Seq)
38	949.5	58.5	930	7	AAC38650 Prokaryot
39	949.5	58.5	933	7	ACA40638 Prokaryot
40	949.5	58.5	110000	4	AA199682-25 Continuation (26 o
41	949.5	58.5	110000	4	AA199682-26 Continuation (26 o
42	949.5	58.5	110000	4	AA199683-25 Continuation (27 o
43	949.5	58.0	930	7	ACA38024 Prokaryot
44	941.5	58.0	1086	7	ACA38210 Prokaryot
45	941.5	58.0	1086	7	ACA38210 Prokaryot

## ALIGNMENTS

### RESULT 1

AAV99906 standard; cDNA; 1362 bp.

AAV99906;

27-SRP-1999 (first entry)

Fragment of cysteine synthase gene.

Biosynthesis; biosynthetic pathway; lysine; threonine; methionine;

cysteine; isoleucine; amino acid; homoserine kinase; aspartic semialdehyde dehydrogenase; diaminopimelate decarboxylase;

cysteine synthase; cystathione beta-lyase; gene expression; screening; inhibition; ss.

Glycine max.

Key Location/Qualifiers

FT CDS 90..1067 /tag= a /product= "Cysteine synthase fragment"

FT FT W09856935-A2.

FT PD 17-DEC-1998.

FT PF 11-JUN-1998; 98WC-US012073.

PR 12-JUN-1997; 97US-0049406P.

PR 12-NOV-1997; 97US-0065385P.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Palco SC, Allen SM, Thorpe CJ;

XX MPI: 1999-080910/07.

XX P-PSDB; AAW81018.

XX New isolated plant amino acid biosynthetic enzyme nucleic acids - which  
PT encode aspartic semialdehyde dehydrogenase, diaminopimelate  
PT decarboxylase, homoserine kinase, cysteine synthase and cystathionine  
PT beta-lyase.

XX Claim 17; Page 57-58; 80pp; English.

XX Organisation of the pathway leading to plant biosynthesis of lysine,  
CC threonine, methionine, cysteine and isoleucine suggests that over-  
CC expression or reduction of expression of genes encoding enzymes involved  
CC in that biosynthetic pathway could be used to alter the level of these  
CC amino acids in human food and animal feed. This may increase the  
CC nutritional quality of human food and animal feed by increasing the  
CC production and accumulation of specific free amino acids. The enzymes  
CC include aspartic semialdehyde dehydrogenase, homoserine kinase,  
CC diaminopimelate decarboxylase, cysteine synthase and cystathionine beta-  
CC lyase. The nucleic acids encoding these enzymes can be used for altering  
CC the level of expression of the enzymes and for evaluating compounds for  
CC their ability to inhibit the enzymes' activity

XX Sequence 1362 BP; 388 A; 243 C; 330 G; 401 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.25e-151 Length: 1362  
Score: 1623.00 Matches: 325  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-931-457A-31 (1-325) x AAV99906 (1-1362)

QY 1 MetAlaValGluArgSerGlyIleAlaValAspValThrGluLeuIleGlyLysThrPro 20

DB 90 ATGGCTGTGAAAGCTCCGGAATGCGAAGATGTTACGGAATGTTGTTAAACCCCA 149

QY 21 LeuValIleuLeuGlyLeuAlaValGluArgSerGlyIleAlaValAspValThrGlu 40

DB 150 TTAGTATATCTAATAAATCACTGCGAGTGTGTGTGCTGCTGCTAAACTGAG 209

QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyIleSerMetIleAlaAspAla 60

DB 210 TTATATGAGGCAATGCTTACGTGAGGAGCAAGATGCGTATGATGATGCTGAGAC 269

QY 61 GluGluLeuGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyLys 80

DB 270 GAAGAGAGAGGAGCTTACCACTGCGAAGAGTGTCTCTATGACCAACAAAGTGTAAAT 329

QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyIleIleLeuLeuAlaIleThrMet 100

DB 330 ACTGGCACTGATTAAGCTTCAATGGCAGCGCGGCTTACAAAGCTCATTAATTAACAG 389

QY 101 ProAlaSerMetSerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120

DB 390 CCGCTCTCTATGAGCTTGAAGAGAGATCACTTATACCTTTGAGAGCTGAGTGTGT 449

QY 121 LeuThrAspProAlaValGlyMetLysGlyAlaValGluIleLysAlaGluIleLeuAla 140

DB 450 CTGACAGATCTCTGTAAGGAGATGAAAGTCTGCTCAGAGGCTGAGAGATATGGCT 509

QY 141 LysThrProAlaValIleLeuGluIlePheGluAsnProAlaAsnPolysValHis 160

DB 510 AAGAGCGCCCAATGCTCAATCACTTCAACAAATTTGAAACCCCTGCCAATCCCAAGTTCAAT 569

QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180

DB 570 TATGAACCACTGCTCCAGATATGAAAGGCTCCGATGCGAAATGATGATTTGTT 629

QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysIleLeuGluAsn 200

DB 630 TCTGGATAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689

QY 201 ProAlaIleLysLeuIleGlyValGluProValGluIleSerProValLeuSerGlyLys 220

DB 690 CCGAATATTAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749

QY 221 ProGlyProHisLysIleGluIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 240

DB 750 CCGTGTCCACACAAATTCAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 809

QY 241 AsnLeuLeuAspGluValValGluIleSerSerAspGluAlaIleGluThrAlaLysLeu 260

DB 810 ATCTCTTGTGATGAGAGTGTGCTCAATATCAAGTATGAGCAATGAAATGCAAGCTT 869

QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

DB 870 CTGGCTTAAAGAGGCTTATTTGTGGATATCTTCGGAGCTGCAAGCTGCTGCT 929

QY 281 PheGluIleAlaValArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300

DB 930 TTTCAGATTGCAAAAGAGCAAGAAATCCCGGAGCTTATTTGCTGCTTTTCCAGC 989

QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAlaGluSer 320

DB 990 TTCCGGGAGAGGTACCTGCTCCCTGCTATTTGAGTCAAGTCAAGAGAGCTGAAAGC 1049

QY 321 MetThrPheGluPro 325

DB 1050 ATGACCTTTGAGCC 1064

RESULT 2

AAQ4450 standard; cDNA to mRNA, 1303 BP.

AAQ4450;

12-OCT-1994 (first entry)

Cysteine synthase.

Cysteine synthase; plant; expression; probe; ss.

Spinacia oleracea.

Key Location/Qualifiers

FT CDS 52..1029

FT /tag= a

PN JP06038770-A.

PD 15-FEB-1994.

XX 05-FEB-1992; 92JP-00020315.

XX 05-FEB-1992; 92JP-00020315.

XX (MITS) MITSUBISHI CORP.

XX (MITU) MITSUBISHI KASEI CORP.

XX MPI: 1994-094834/12.

XX P-PSDB; AAR49830.

XX Novel gene coding cysteine synthase - used to increase the cysteine

XX content of an agricultural plant.

XX Claim 2; Page 4-6; 6pp; Japanese.

XX The cysteine content in an agricultural product can be increased by

XX expressing the cysteine synthase in a plant. Probes V822 and V812 used in







```

PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145151P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153589P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154038P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156589P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-015656P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157535P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,276-127
Score: 1384.00
Percent Similarity: 92.79%
Best Local Similarity: 83.39%
Query Match: 85.27%
DB: 3
Gaps: 0
US-09-931-457a-31 (1-325) x AAC42360 (1-1120)

QY 6 SerGlyIleAlaValThrGluLeuIleGlyIleValThrProLeuValIleLeuAsn 25
158 TCAGAAATTCCTAAGATGTCAGTCAATTCATTCAGTGAACACCTTCATTCGTTAATTCAGAC 217
Db 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45
QY 218 AATGTCGTAAGAGATGTCGTCGTCCTAAGCTTGAAGATGTAACCGTTC 277
Db 46 SerSerValIysAspArgIleGlyIleValSerMetIleAlaAspAlaGluGluLysGlyLeu 65
278 TCAGAGTCGTAAGAGATGTCGTTTACATTCATTCGACAAAGAAAGAGCGTCTT 337
QY 66 IleThrProGlyIysSerValLeuIleGluProThrSerGlyValThrGlyIleGlyLeu 85
Db 338 ATCAACCGAGAGAGATGTCGTCATTCATTCAGCCAAAGAGAGAACTGCGAGTGGGTTA 397
QY 86 AlaPheMetAlaAlaAlaArgGlyIleValSerLeuIleIleThrMetProAlaSerMetSer 105

```

Db 398 GCATTCAGCGGACGCTGCAAGGCTACAGCTTATATTACATGCGAGCTCTATGAGT 457  
 QY 106 Lcugluarargrilelleleuenualepheglialagluenuvalleuthrapproala 125  
 Db 458 ACTGAGAGAGAAATCATCTTCTTAGCTTTGAGGTGATGTTTAACTGACCAAGCT 517  
 QY 126 LysGlyMetLysGlyValAlaGlnLysValAlaGlnLysValAlaGlnLysValAlaGln 145  
 Db 518 AAGGCGATGAAAGAGCTATCGCAAGCGGAAAGATTTGGCGAAGAAACCAACATGCT 577  
 QY 146 TTTTleuenginginephgluasnpProAlaaspProLysValHisTyrGluThrGly 165  
 Db 578 TACATGCTCAGCAGTTTGAGAAACCTGCAACCTTAAATCCATGAGACTACGGGA 637  
 QY 166 ProGluThrLysGlySerAspGlyLysValAlaAspAlaPheValSerGlyIleGlyThr 185  
 Db 638 CTTGAGATATGAAAGAGCTGCTGGCAAAATGATGGCTTTGTTCTGGAATGGTACT 657  
 QY 186 GlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlnLysAsnProAsnIleLysLeu 205  
 Db 698 GGTGGTACCATTCAGGCTGCGGAGATCTTAAAGAACAGAACGCAACGTCAGCTG 757  
 QY 206 IlleglyValGluProValGluSerProValLeuSerGlyLysProGlyProAla 225  
 Db 758 TATGAGATGAGAGCCAGTTGAAAGGTCTATCTATCCGCTGGAGAGCCAGTCTCAACAG 817  
 QY 226 IlleginglyIleGlyAlaGlyPheIleProGlyValleuGluValAsnLeuAspGly 245  
 Db 818 ATTCAAGGATATGAGAGCTGTTTATACAGAGTATGATGATGATCTTATTTGACGA 877  
 QY 246 ValValGlnLysSerSerAspGluAlaIleGluThrAlaLysLeuAlaLeuLysGln 265  
 Db 878 GTTGTTCAGGTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 937  
 QY 266 GluLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlnIleAlaLys 285  
 Db 938 GGTCTTCTTGTGGATATCATTCGCTGCAAGCTGCTGCAAGATTAACCTTGCACAG 997  
 QY 286 ArgProGluAsnAlaGlyLysIleValAlaValaPheProSerPheGlyGluThr 305  
 Db 998 AGCCCAAGAAACCGCTGGAGAGCTATTTGCGCGATATTCCTCGAGTTTCGGGAGAGTAA 1057  
 QY 306 LeuSerSerValleuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
 Db 1058 CTATCAACGGTACTTTCGATGCGACAGAAAGAAAGAGAGAGAGAGAGAGAGAGAG 1114  
 RESULT 5  
 ABZ13088  
 ID ABZ13088 standard; DNA; 966 BP.  
 AC ABZ13088;  
 DT 21-JAN-2003 (first entry)  
 XX Arabidopsis thaliana stress regulated gene SEQ ID NO 893.  
 XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 OS Arabidopsis thaliana.  
 PN WO200216655-A2.  
 PD 28-FEB-2002.  
 PF 24-AUG-2001; 2001MO-US026685.  
 PR 24-AUG-2000; 2000US-0227866P.  
 PR 26-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.  
 PA (SCRI) SCRIPS RES INST.  
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;  
 DR WPI; 2002-304127/34.  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 PS Claim 14; SEQ ID NO 893; 577bp + Sequence Listing; English.  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ1574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 SQ Sequence 966 BP; 280 A; 157 C; 261 G; 268 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,02e-126 Length: 966  
 Score: 1374.00 Matches: 267  
 Percent Similarity: 91.08% Conservative: 29  
 Best Local Similarity: 82.15% Mismatches: 25  
 Query Match: 84.66% Gaps: 4  
 DB: 6 Gaps: 1  
 US-09-931-457A-31 (1-325) x ABZ13088 (1-966)  
 QY 1 MetAlaValAlaGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
 Db 1 ATGGCTTACGAGTGCCTCTTAAATCCCTAAAGATGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 60  
 QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaLysLeuGlu 40  
 Db 61 TTAGGTATCTTAACCAATAGTACTTAAGATTTGTTGCTCAGTCTGCTAGCTTGG 120  
 QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 121 ATGATGGAACCAATGTTCTAGCGTCAAGCAAGATGGTTATATGATGATGATGATGATGAT 180  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValleuIleGluProThrSerGlyAsn 80  
 Db 181 GAAGCTAAAGACTTATTAACCGGAGAGAGATGTTGATGATGATGATGATGATGATGATGAT 240  
 QY 81 ThrGlyIleGlyLeuAlaPheSerAlaAlaAlaArgIlyTyrLysLeuIleIleThrMet 100  
 Db 241 ACTGAGATGTTGTTTACATTTACTGAGCTGCAAGCGCTATTAAGCTTGTATTAACAATG 300  
 QY 101 ProAlaSerMetSerLeuGluArgGlyIleleleuenualepheglialagluenuval 120  
 Db 301 CCAGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 QY 121 LeuThrAspProAlaLysGlyMetLysGlyValAlaGlnLysValAlaGlnLysValAla 140  
 Db 361 CTTCATGATCCCGCTAAAGTATGAAAGGTGCTGTTGCCAAGCGGAGAGATTTTGGCG 420  
 QY 141 LysThrProAsnAlaLysIleleuenginginephgluasnpProAlaaspProLysValHis 160  
 Db 421 AAACACCTAATAGTATATGCTTCAACAGTTTGGAATCCGCTAATCAAGATCCAT 480  
 QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAlaPheVal 180  
 Db 481 TATGAGATCTACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsn 200  
 Db 541 TCTGATATGTAATCGT-----GGTCTGGGAGATATCTCAAGAAACGAAC 588

Qy	201	ProAsnIleuysIeuIleGlyValGluProValGluSerProValIleuSerGlyValys	220
Db	589	ACAAACATTAAGCTGTAATGCTGCAGACCTCTGTAAAGCCATATTCTCCGGCGGAAG	648
Qy	221	ProGlyProHisIleuIleGlnGlyIleGlyValaGlyPheIleProGlyValIleuGluVal	240
Db	649	CCAGTCCCAACAATAATTCAGAGTAAAGAGTGTCTGTTTATATCCAGCATTTTGAAGTT	708
Qy	241	AsnIleuIeuAspGluValValGlnIleSerSerAspGluValIleGluThrAlaIysIeu	260
Db	709	GATTTATATAGATAGATGATGTTTCAGGTTTCAAGTGAAGAGATCATTCAGATGCGCAAGCTT	768
Qy	261	IeuuIleuIeuysGlnGlyIeuPheValaGlyIleSerSerGlyValaIalaIalaIalaIa	280
Db	769	CTTGCTAGGGAAGAAGGCTCTTGTAGTGGAAATCTCATCTGACAGACAGCTACTGACAGA	828
Qy	281	PheGlnIleAlaIalysArgProGluAsnAlaGlyIysIeuIleValaIalaValaPheProSer	300
Db	829	ATCAAACTTGCAAAGAGGCCACAGAAATATCTGGAGAAAGCTTATATATGCGCGTGTCCAGT	888
Qy	301	PheGlyIuIuArgIylIeuSerSerValIleuPheGluSerValATGArgGluIalaGluSer	320
Db	889	TTCCGGGGAAGAAATCATCTATCGACTGTACTGTTATGACAGCAAGAAAGAGGCGAGAACT	948
Qy	321	MetThrPheGluPro	325
Db	949	ATGACCTTTGAGCCC	963

## RESULT 6

ID AD868423 standard; DNA; 966 BP.  
XX  
XX  
AC AD868423;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Arabidopsis thaliana gene, SEQ ID 683.  
DE  
XX  
XX Plant; bacterial infection; fungal infection; viral infection; ds.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX WO2003000898-A1.  
XX  
XX  
XX 03-JAN-2003.  
XX  
XX  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX  
XX (SYGN ) SYNGENETIC PARTICIPATIONS AG.  
XX  
XX  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,  
PI Pi Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G,  
XX  
XX WPI; 2003-175290/17.  
XX

PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
DS Claim 6, SEQ ID NO 683; 899pp; English.

CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to  
illustrate the invention.  
CC  
CC  
XX  
XX  
SQ Sequence 966 BP; 280 A; 157 C; 261 G; 268 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,02e+126	length:	966
Score:	1374.00	Matches:	267
Percent Similarity:	91.08%	Conservative:	29
Best Local Similarity:	82.15%	Mismatches:	25
Query Match:	84.66%	Indels:	4
DB:	7	Gaps:	1

US-09-931-457A-31 (1-325) X ADA68423 (1-966)

QY MetAlValGlunbSergIylIeAlaIysAspValThrGluLeuIleGlyLySThrPro 20  
 Db 1 ATGGCTTCACTGGCTCCTAAATTCCTAAAGATGTAATGCAATCAATATAGGGAACAATCCA 60  
 QY 21 LeuValIyTrIleuAsnIyLeuAlaAspGlyCysValAlaArgValAlaAlaIySLeuGlu 40  
 Db 61 TTGTGTACTTAAACCAATAGCTTAAAGATGTGTGTGTCAATGTGTGTCAACTTGAG 120  
 QY 41 LeuMetGluProCysSerSerValIyAspArgIleGlyIyrsMetIleAlaAspAla 60  
 Db 121 ATATGTGAACATGTTCTAGCTCAAAAGACAGATGTGTTATAGTATCATTCGTCAAGCA 180  
 QY 61 GluGluIySgIyLeuIleThrProGlyIySerValLeuIleGluProThrSergIyAsn 80  
 Db 181 GAAGCTAAAGACCTATTATTAACCGGAGACAGTGTGTGATGTAGCCAAACAAGGGGAC 240  
 QY 81 ThrGlyIyIleGlyIyLeuAlaPheMetAlaAlaIaArgIyIyTrIySLeuIleIleThMet 100  
 Db 241 ACTGGATGGTGTAGCATTTACTCAAGTGCMAAGGGCTATTAAGCTTTGTTATACATG 300  
 QY 101 ProAlaSerMetSerIeuGluIuArgIleIleLeuLeuAlaPheGlyIyAlaGluLeuVal 120  
 Db 301 CCAGCTTCATATGATATAGAGAGAAATTCATTCCTTACGTTTGGACGTGAGTTGATT 360  
 QY 121 LeuThrAspProAlaIySgIyMetIySgIyAlaValGluIySAlaGluGluIleLeuAla 140  
 Db 361 CTACTGATTCGGCTTAAGGTATGAAAGGTGTGTGTCMAAGCGCGAGAGATTTGGCG 420  
 QY 141 LysThrProAsnAlaTrIleLeuGluIyIlePheGluIyIleAsnProIyIyAsnProIyIyHis 160  
 Db 421 AAAACACCTAATGGTATATATGCTTCAACAGTTTGAGAAATCTGCTTAATCCAAAGATCAT 480  
 QY 161 TyrGluIyThrGlyIyProGluIleTrpIySgIySerAspGlyIyIyIleAspAlaPheAl 180  
 Db 481 TATAGAGCATACGACCTGATGATAGAAAGGTTCTGTGTGGAATAAGTCAATGCTTTGTT 540  
 QY 181 SerGlyIleGlyIyThrIyGlyIyThrIleThrGlyIyAlaGlyIySyrIyLeuIySgIyGluIyAsn 200  
 Db 541 TCTGGATATGTAATCGT-----GGTCCGAGAGTATCCAGAGAAACGAAAC 588  
 QY 201 ProAsnIleIySLeuIleGlyIyAlaGluProIyAlaGluSerProValIyLeuSergIyGlyIyS 220  
 Db 589 ACAACATTAAGCTGATATGATGTGCAAGCTGTGTAAGCCATATTCGTCCGGTGGAAAG 648  
 QY 221 ProGlyIyProIyAlaIyIleGluIyIleGlyIyAlaGlyPheIleProGlyIyValLeuGluVal 240  
 Db 649 CCAGGCTCCCACAAAATTCAGAGTATAGTCTCGTGTATTATCCAGGACATTTGGAGTT 708  
 QY 241 AsnIleuAsnAspGluValIyAlaGluIleSerAspGluAlaIleGluTrpAlaIyLeu 260  
 Db 709 GATCTTATATGAATGATGTTCTCAAGTTCAAGTAAAGATTCATTGACATGCAAGCTT 768  
 QY 261 LeuAlaIleuIySgIyGluIyIyLeuPheValIyIleSerSergIyAlaAlaAlaAlaAla 280  
 Db 769 CTGTGCTAGAGAAAGAGGCTCTGTATGTGGAATTCATCTCGAGCGACAGCTACTGCGACGA 828  
 QY 281 PheGluIleAlaIyArgIyProGluAsnAlaGlyIySyrIyLeuIyValAlaValPheProSer 300



XX Oryza sativa.  
 OS WO2003000898-A1.  
 PN 03-JAN-2003.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,  
 PI Karaghi F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;  
 DR WPI; 2003-175290/17.  
 XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX Claim 6; SEQ ID NO 3760; 899pp; English.  
 XX The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX Sequence 5259 BP; 1631 A; 1084 C; 1282 G; 1261 T; 0 U; 1 Other;  
 SO  
 Alignment Scores:  
 Pred. No.: 2,966-124 Length: 5259  
 Score: 1359.00 Matches: 260  
 Percent Similarity: 89.72% Conservative: 28  
 Best Local Similarity: 81.00% Mismatches: 33  
 Query Match: 83.73% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-931-457A-31 (1-325) x ADA70437 (1-5259)  
 QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyIleThrPro 20  
 DB 1 ATGGCGGAATCATGCGCATCGATTCCGAGTATGAGTGCACATCGCGCAATATCTCC 60  
 QY 21 LeuValIleuAsnIleuAlaAspGlyCysValAlaArgValAlaAlaIleuGlu 40  
 DB 61 CTGGGTACTCGTGAACAGAGTGGTGGATGATGCGCCAAATCGCTGCCAAGCTTAG 120  
 QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyIleSerMetIleAlaAspAla 60  
 DB 121 ATCATGAGCCCTGCTCCAGTGTCAAGATAGATGTTATAGCATATATTCAGATGCA 180  
 QY 61 GluGluIysGlyIleuIleThrProGlyIysSerValIleuIleGluProThrSerGlyAsn 80  
 DB 181 GAAGAGAGAGGCTCATCTACCAAGAAAGAGTGTGATTGATTGAGCCCTACATGAGAAAT 240  
 QY 81 ThrGlyIleGlyIleuAlaPheMetAlaAlaAlaArgIleValIleuIleIleThrMet 100  
 DB 241 ACTGATATGCTCTGCTTCAATGCTCTGCTGCAAGGCTTAACTATATCTGCAATG 300  
 QY 101 ProAlaSerMetSerLeuGluArgIleIleIleuAlaPheGlyAlaGluLeuVal 120  
 DB 301 CCAAGATCATGAGATGAGAAAGAGATCATTTTAAGGCTTCGTCGCGAGTTGCTC 360  
 QY 121 LeuThrAspProAlaIleuGlyMetCysGlyAlaValGluIysAlaGluIleuAla 140

DB 361 CTATCCCATCATCTTCGGGCAATGAAGTGCATCAAGAGGAGATGAGCTGCAGCA 420  
 QY 141 LysThrProAlaMetAlaIleuGluGlnPheGluAsnProAlaAsnProIysValHis 160  
 DB 421 AAGATGCCCTAATCTTTCATCTCCAGCAAGTTGAGAACCTTGAACCCCAAAATATCCAC 480  
 QY 161 TyrGluThrThrGlyProGluIleThrIlePheGlySerAspGlyIleAspAlaPheVal 180  
 DB 481 TATAGCAACTGACCTGATGATTTGGAAGGCCACTCTGCGAAGATGATATCTTCTC 540  
 QY 181 SerGlyIleGlyThrGlyIleThrIleGlyAlaGlyIleValIleuIysGluGlnAsn 200  
 DB 541 TCTGCAATGGAAACCGGTGTACAGTAAACCGAAACAGAAAGTACCTCAAGAGACAGAAC 600  
 QY 201 ProAsnIleIysIleuIleGlyValGluProValGluSerProValIleuSerGlyIlys 220  
 DB 601 CTGAGATCAAGATCTATGCTGTGAACCAAGAAAGTCAATTTATCTGAGAGAGA 660  
 QY 221 ProGlyProHisIysIleGlnGlyIleGlyAlaGlyIlePheIleProGlyValIleuVal 240  
 DB 661 CTGTGTCACACAAATCCAAAGGATAGTGTGCTGTTTCTCTGCTGATCTGATGTC 720  
 QY 241 AsnLeuIleAspGluValIleGlnIleSerSerAspGluAlaIleGluThrAlaIleu 260  
 DB 721 AATCTCTGATGAAGTGTCCAGTCTCAAGTATGAAGCATCAAGATGCGAGACAG 780  
 QY 261 LeuAlaIleuIysGluIleuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
 DB 781 CTACGCTGAAGAAGAGGCTGCTGTTGGATCTCATCTGCTGCTGCGAGTCGCGCC 840  
 QY 281 PheGlnIleAlaIysAspProGluAsnAlaGlyIleuIleValAlaValPheProSer 300  
 DB 841 ATCAGGTCCGCGACCGCCCGCAAGAGGAGGCTCTAGTGTGCTCTCCAGC 900  
 QY 301 PheGlyGluArgIleuSerSerValIleuPheGluSerValAlaArgArgIleuGluSer 320  
 DB 901 TTCGTAGCGGCTACCTGCTGCTCTCTCTCTTGGATCATCAAGAGAGGCGGAAAC 960  
 QY 321 Met 321  
 DB 961 ATG 963  
 RESULT 9  
 AAQ74413 standard, cDNA to mRNA, 1483 BP.  
 XX AAQ74413;  
 AC 12-JUN-1995 (first entry)  
 DT 12-JUN-1995 (first entry)  
 XX Spinach cysteine synthase.  
 DE Spinach cysteine synthase.  
 XX Cysteine; spinach; spinacia oleracea; pUC19; M13mp18; plant; nutrition;  
 KW feed value; pharmaceutical; food additive; cosmetic; E.coli; ds.  
 XX Spinacia oleracea.  
 OS  
 Key Location/Qualifiers  
 FT CDS 62..1213 /product= "cysteine synthetase"  
 FT /tag= a  
 FN JP06245773-A.  
 PD 06-SEP-1994.  
 XX 26-FEB-1993; 93JP-00038527.  
 XX 26-FEB-1993; 93JP-00038527.  
 XX 26-FEB-1993; 93JP-00038527.  
 XX (MITS) MITSUBISHI CORP.  
 PA (METU) MITSUBISHI KASEI CORP.



Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rng

Page 11

PR 19-MAY-1999; 99US-0139441P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140351P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147392P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149262P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150864P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153079P.  
PR 13-SEP-1999; 99US-0153786P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160983P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.





QY 16 IIEGLVYLeuThrProLeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgVal 35  
 Db 361 ATTGGGAAACTCTCATGGGTGACTTGAACAGTATGCAAAAGTTGTGTACCAATAT 420  
 QY 36 AlaAlaLysLeuGluLeuMetGluProCysSerSerValLysAspArgLleGlyTyrSer 55  
 Db 421 GCTGCCAAACTCGAGATCATGAGCAGCATGTTCAAGTCAAGATGAGATTGGGTACAGT 480  
 QY 56 MetIleAlaAspAlaGluGluLysGlyLeuLleThrProGlyLysSerValLeuLleGly 75  
 Db 481 ATGGTACTGATGCTAGCAAAAAGATTATATATCACTCGAAAGAGTGTCTTGGGAA 540  
 QY 76 ProThrSerGlyAsnThrGlyLleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLys 95  
 Db 541 CTTACAGAGTGAACACCGGTATTTGGCTTACGATTCATGCTGCTTCAAGGGCTATAGA 600  
 QY 96 LeuLleLleThrMetProAlaSerMetSerLeuGluArgArgLleLleLeuLeuAlaPhe 115  
 Db 601 CTTATCTTGCACATGCCAGCATCCAGATGAGAAAGAGGTTCTTTTAAAGGCATT 660  
 QY 116 GlyAlaGluLeuValLeuThrAspProAlaLysGlyMetLysGlyAlaAlaGluLysAla 135  
 Db 661 GAGAGCTGACCTGTATTAACAATCTCTGTAAGGATATCATGAGCGCTTCAAGAGCT 720  
 QY 136 GluGluLleLeuAlaLysThrProAsnAlaTyrLleLeuGluGlnPheGluAsnProAla 155  
 Db 721 GAAGAGATCTTGAAAAACACCTCTGATGCATCATGCTCCAAAGTTGACAACTCTGCA 780  
 QY 156 AspProLysValHisTyrGluThrThrGluProGluLleThrLysGlySerAspGlyLys 175  
 Db 781 AATCCAAAGATTCATTACAGAGCTACTGTGCTCTGAGATTTGGAGATGATCAAAAGGCAG 840  
 QY 176 IleAspAlaPheValSerGlyLleGlyThrGlyGlyThrLleThrGlyAlaGlyLysTyr 195  
 Db 841 GTGAGATCTTGTGTCAGAGATTTGGAAGTGTGGAATCATCATGAGTGTGTGCTGATTT 900  
 QY 196 LeuLysGluGlnAsnProAsnLleLysLeuLleGlyValGluProValGluSerProVal 215  
 Db 901 ATTTAAAGAGAAAAATCCCAAAAGCGATTATGGTGTAGAAACCAACGAAAGCGATATA 960  
 QY 216 LeuSerGlyGlyLysProGlyProHisLysLleGlnGlyLleGlyAlaGlyPheLlePro 235  
 Db 961 CTTTCTGTGGGAAACCGGAGCTCATCAAGATTCAGAAATTTGGAGCTGGATTATACCT 1020  
 QY 236 GlyValLleGluValAsnLeuLysAspGluValAlaGlnLleSerSerAspGluAlaIle 255  
 Db 1021 AAGAAATTTGATCAGAAATTTATGATGATGATCATAGCATTTCAAGTGAAGAACTATA 1080  
 QY 256 GluThrAlaLysLeuLeuAlaLeuLysGluGlyLysPheValGlyLleSerSerGlyAla 275  
 Db 1081 GAAACCGCAACCACTACTCTTTAAGAGAGCTTGATGATGTTGTATCATCTGAGACT 1140  
 QY 276 AlaAlaAlaAlaAlaPheGlnLleAlaLysArgProGluAsnAlaGlyLysLeuLleVal 295  
 Db 1141 GAGAGCTGCTGCTGCAATCAAGGTGCTAAGAGACCTGAGAAAGCTCGAAACTCATGCT 1200  
 QY 296 -----AlaValPheProSerPheGlyGlu 303  
 Db 1201 GTAACGTATTAATGAATCTATGAATGCGGTTTTCAGGTGTGTTTCCCGAGCTTGGGAG 1260  
 QY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMet 322  
 Db 1261 AGATATCATATGACCCCGCTGTTCCAGTTCATCCGGGAAGAGGTGAGAAATATG 1314  
 RESULT 12  
 AAC40774 standard; DNA; 1224 BP.  
 AAC40774;  
 AC 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 29501.

XX Hybridisation assay; Genetic mapping; Gene expression control;  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 KM Arabidopsis thaliana.  
 PN EP1033405-A2.  
 PD 06-SEP-2000.  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-013548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 23-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132566P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138647P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139482P.  
 PR 18-JUN-1999; 99US-0139484P.  
 PR 18-JUN-1999; 99US-0139485P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P  
PR 23-JUN-1999; 99US-0140354P  
PR 23-JUN-1999; 99US-0140685P  
PR 28-JUN-1999; 99US-0140823P  
PR 29-JUN-1999; 99US-0140981P  
PR 30-JUN-1999; 99US-0141287P  
PR 01-JUL-1999; 99US-0141842P  
PR 01-JUL-1999; 99US-0142154P  
PR 02-JUL-1999; 99US-0142055P  
PR 06-JUL-1999; 99US-0142350P  
PR 09-JUL-1999; 99US-0142803P  
PR 12-JUL-1999; 99US-0142977P  
PR 13-JUL-1999; 99US-0143542P  
PR 14-JUL-1999; 99US-0143624P  
PR 15-JUL-1999; 99US-0144005P  
PR 16-JUL-1999; 99US-0144085P  
PR 16-JUL-1999; 99US-0144086P  
PR 19-JUL-1999; 99US-0144325P  
PR 19-JUL-1999; 99US-0144331P  
PR 19-JUL-1999; 99US-0144332P  
PR 19-JUL-1999; 99US-0144333P  
PR 19-JUL-1999; 99US-0144334P  
PR 19-JUL-1999; 99US-0144335P  
PR 20-JUL-1999; 99US-0144352P  
PR 20-JUL-1999; 99US-0144632P  
PR 20-JUL-1999; 99US-0144884P  
PR 21-JUL-1999; 99US-0144814P  
PR 21-JUL-1999; 99US-0145086P  
PR 21-JUL-1999; 99US-0145087P  
PR 22-JUL-1999; 99US-0145087P  
PR 22-JUL-1999; 99US-0145087P  
PR 22-JUL-1999; 99US-0145089P  
PR 22-JUL-1999; 99US-0145192P  
PR 23-JUL-1999; 99US-0145145P  
PR 23-JUL-1999; 99US-0145218P  
PR 23-JUL-1999; 99US-0145247P  
PR 26-JUL-1999; 99US-0145276P  
PR 27-JUL-1999; 99US-0145913P  
PR 27-JUL-1999; 99US-0145918P  
PR 27-JUL-1999; 99US-0145919P  
PR 28-JUL-1999; 99US-0145951P  
PR 02-AUG-1999; 99US-0146386P  
PR 02-AUG-1999; 99US-0146388P  
PR 02-AUG-1999; 99US-0146389P  
PR 03-AUG-1999; 99US-0147038P  
PR 04-AUG-1999; 99US-0147204P  
PR 04-AUG-1999; 99US-0147302P  
PR 05-AUG-1999; 99US-0147192P  
PR 05-AUG-1999; 99US-0147260P  
PR 06-AUG-1999; 99US-0147303P  
PR 06-AUG-1999; 99US-0147416P  
PR 09-AUG-1999; 99US-0147493P  
PR 09-AUG-1999; 99US-0147935P  
PR 10-AUG-1999; 99US-0148171P  
PR 11-AUG-1999; 99US-0148319P  
PR 12-AUG-1999; 99US-0148341P  
PR 13-AUG-1999; 99US-0148565P  
PR 13-AUG-1999; 99US-0148684P  
PR 16-AUG-1999; 99US-0149368P  
PR 17-AUG-1999; 99US-0149175P  
PR 18-AUG-1999; 99US-0149426P  
PR 20-AUG-1999; 99US-0149722P  
PR 20-AUG-1999; 99US-0149723P  
PR 20-AUG-1999; 99US-0149929P  
PR 23-AUG-1999; 99US-0149902P  
PR 23-AUG-1999; 99US-0149930P  
PR 25-AUG-1999; 99US-0150566P  
PR 26-AUG-1999; 99US-0150884P  
PR 27-AUG-1999; 99US-0151065P  
PR 27-AUG-1999; 99US-0151066P  
PR 27-AUG-1999; 99US-0151080P  
PR 30-AUG-1999; 99US-0151303P

PR 31-AUG-1999; 99US-0151438P  
PR 01-SEP-1999; 99US-0151930P  
PR 07-SEP-1999; 99US-0152363P  
PR 10-SEP-1999; 99US-0153070P  
PR 13-SEP-1999; 99US-0153758P  
PR 15-SEP-1999; 99US-0154018P  
PR 16-SEP-1999; 99US-0154039P  
PR 20-SEP-1999; 99US-0154779P  
PR 22-SEP-1999; 99US-0155139P  
PR 23-SEP-1999; 99US-0155486P  
PR 24-SEP-1999; 99US-0155659P  
PR 28-SEP-1999; 99US-0156458P  
PR 29-SEP-1999; 99US-0156596P  
PR 04-OCT-1999; 99US-0157117P  
PR 05-OCT-1999; 99US-0157753P  
PR 06-OCT-1999; 99US-0157865P  
PR 07-OCT-1999; 99US-0158029P  
PR 08-OCT-1999; 99US-0158232P  
PR 12-OCT-1999; 99US-0158369P  
PR 13-OCT-1999; 99US-0159293P  
PR 13-OCT-1999; 99US-0159294P  
PR 13-OCT-1999; 99US-0159295P  
PR 14-OCT-1999; 99US-0159329P  
PR 14-OCT-1999; 99US-0159330P  
PR 14-OCT-1999; 99US-0159331P  
PR 14-OCT-1999; 99US-0159637P  
PR 14-OCT-1999; 99US-0159638P  
PR 18-OCT-1999; 99US-0159584P  
PR 21-OCT-1999; 99US-0160741P  
PR 21-OCT-1999; 99US-0160767P  
PR 21-OCT-1999; 99US-0160768P  
PR 21-OCT-1999; 99US-0160770P  
PR 21-OCT-1999; 99US-0160814P  
PR 21-OCT-1999; 99US-0160815P  
PR 22-OCT-1999; 99US-0160980P  
PR 22-OCT-1999; 99US-0160981P  
PR 22-OCT-1999; 99US-0160982P  
PR 25-OCT-1999; 99US-0161404P  
PR 25-OCT-1999; 99US-0161405P  
PR 25-OCT-1999; 99US-0161406P  
PR 25-OCT-1999; 99US-0161406P  
PR 26-OCT-1999; 99US-0161359P  
PR 26-OCT-1999; 99US-0161360P  
PR 26-OCT-1999; 99US-0161361P  
PR 28-OCT-1999; 99US-0161920P  
PR 28-OCT-1999; 99US-0161992P  
PR 28-OCT-1999; 99US-0161993P  
PR 29-OCT-1999; 99US-0162142P

Alignment Scores:  
Pred. No.: 6,33e-109  
Score: 1196.00  
Percent Similarity: 84.74%  
Best Local Similarity: 71.34%  
Query Match: 73.63%  
DB: 3  
Gaps: 0

US-09-931-457a-31 (1-325) x AAC40774 (1-1224)

QY 4 GUARGSERGLYLALALYSPVALTHRGULEUIGLYLYSTHPRULEVALTYR 23  
DB 51 GARGGCTTCTATCAAGACATGCTACTGATTTGTTGTAACACCAATGCTAT 110  
QY 24 LEUASUPLYLEUALASPGLYCYVALALARGVALALALALYSLUGULEUWECGU 43  
DB 111 CTGACCAANGTGTGATGCTGCTGCTGCTATGCTGCAACCTTGATGATGAG 170  
QY 44 PROCVSSERVALYASPARGLIEGLYRISERMETLEALASPAIAGIUGLYS 63  
DB 171 CTTGTCTTCAAGCTAAGACAGAAATCCGCTATGATGATCAAGACGACAGAAA 230  
QY 64 GYLEULIETHPRGGLYLYSERVALLEULEIGIUPROTHSERGLYASANTHGLY 83  
DB 231 GGATGATTATCCCGGAAAGGTACATTGATAGGCCAAGCTGTTAACAACCGGAT 290









```

PR      25-OCT-1999;      99US-0161406P.
PR      26-OCT-1999;      99US-0161359P.
PR      26-OCT-1999;      99US-0161350P.
PR      26-OCT-1999;      99US-0161361P.
PR      28-OCT-1999;      99US-0161920P.
PR      28-OCT-1999;      99US-0161992P.
PR      28-OCT-1999;      99US-0161993P.
PR      29-OCT-1999;      99US-0162142P.

Alignment Scores:
Pred. No.:      1,972-108      Length:      1423
Score:      1192.00      Matches:      227
Percent Similarity:      84.74%      Conservative:      45
Best Local Similarity:      70.72%      Mismatches:      49
Query Match:      73.44%      Indels:      0
DB:      3      Gaps:      0

```

US-09-921-457A-31 (1-325) x AAC41206 (1-1423)

QY	4	GLuNRGserGIYllleAlaIyAspValThhGluLeuIleGIYlYsThrProLeuValTyr	23
Db	246	GATGGGCTCTGATCAAGAACGAAGTCTCTGAATTTGGTGTAAACACCAATGGTGTAT	305
QY	24	LeuAsnIyLeuAlaAspGIYCySValAlaArgValAlaIaIySLeuGIuLeuMerGIu	43
Db	306	CTGAACAAATGTTGTGATGGTGGCTGGTCTGGTACCGTCAAGCTTGAATGATGGAG	365
QY	44	ProCysSerSerValIyAspArgIleGIYTySereMetIleAlaAspAlaGIuGIuIyS	63
Db	366	CCTGGTCTACCGCTCAAGACAGATGCCGATGTATGATGATCAAGAAAGTCAAGAGAAAA	425
QY	64	GIYLeuIleTrrProGIYLySerSerValLeuIleGIuProThrSerGIYAspThrGIYle	83
Db	426	GGATTGATTAATCTCCGGAAAGATGACATGATTAAGCCCACTGGCGGTACACCGGATT	485
QY	84	GIYLeuAlaPheMetAlaAlaAlaArgGIYTyIyIySLeuIleIleThrMetProAlaSer	103
Db	486	GGTTTAGCTTCACATGGAGAGCTGCACAGGGCTATAAGGATCCGTGTATCCTTCAACT	545
QY	104	MetSerLeuGIuATGArgIleIleLeuLeuAlaPheGIYAlaGIuLeuValLeuThzAsp	123
Db	546	ATGACCTTAGAGAGAGATCAATCTTGAAGGSCACTAGTCCAGACCTTCATCTCCGAC	605
QY	124	ProAlaIyArgIleMetLySGIYAlaIyAlaGIuIyAlaGIuIyIleAlaIySThrPro	143
Db	606	CAGCGCATAGCGCTTAAAGAAATGTGGAGAAAACTGAAGGATTTTAAAGCAAACTCCT	665
QY	144	AsnAlaTyIleLeuGIuInPheGIuIyAsnProAlaAspProLySValHISTyGIuThr	163
Db	666	GGTGGTTCATTCACACAAATTTGAATCTGCACAAACCCGCAATTCATTACCGAAC	725
QY	164	ThrGIYProGIuIleTrrIyPlySArgIySerAspGIYIySleAspAlaPheValSerGIYle	183
Db	726	AACGGACCGGAAATATGGAGAGATTACACCGCGGAAAGTAGATATATTTGGCTGGCCCTA	785
QY	184	GIYThrGIYGIYThrIleThrGIYAlaGIYIySyrTyIeIySGIuInAspProAsnIle	203
Db	786	GGAGCTGATGAACTGCTACTGAGTATGAGGAAAGTCTCCAGAGACCAAGCAAGACATC	845
QY	204	IySLeuIleGIYAlaGIuProValGIuSerProValLeuSerGIYGIYIyProGIYPro	223
Db	846	AAAGTTTGATGGTGGAGAACAGTATGAATGCCGATCTTAAGCGGAGGTCAACCAAGTCCA	905
QY	224	HISySleGIeGIYIleGIYAlaGIYHelleProGIYValLeuGIuValAsnLeuIyS	243
Db	906	CATTGATTCAGGGAATTTGGCTCTGGTATTCGTGCCATTTGCATTTGAATCAATGTT	965
QY	244	AspGIuValAlaGIuIleSerSerAspGIuAlaIleGIuThrAlaIySLeuAlaIyS	263
Db	966	GATGAATATTATCAAGTGGCAGGTGAAGAAGGCTATTGAACAAGCCAACTTCTTCCCTC	1022
QY	264	LySGIuGIYLeuPheValGIYIleSerSerGIYAlaAlaAlaAlaAlaPheGIuIle	283

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rng

Page 20

Db 1026 AAAGAGGATTACTGTGGGAAATATCCTGTGAGCCGCGACGCGCTGCGTTAAAGTT 1085  
QY 284 AAlAsArpProGluAsnAlAGLYLeuIleValAlaValPheProSerPheGlyGlu 303  
Db 1086 GCANAGCGGCTGAAAAAGCGGGAACCTCTTGTGTGTTTCTTCTAGTGAGAGAA 1145  
QY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPhe 323  
Db 1146 CGTATTTATGAGCTAACTGTGATTGATTGATATGATGAGCAGAAATTGCTATT 1205  
QY 324 Glu 324  
Db 1206 GAA 1208

Search completed: May 21, 2004, 22:24:50  
Job time : 540 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 21, 2004, 20:35:52 ; Search time 4694 Seconds  
(without alignments)  
3000.956 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVESGIAQVTELGKTP.....LSSVLFESVRRAESMTPEP 325

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEP=frame+ p2n.model -DEV=xlh  
-O=csn2.1/USPTO.spool/US09931457.rnat 18052004.121729.20297/app.query.fasta\_1.519  
-DB=csn2.1/USPTO.spool-super+p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-INIT=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -NODE-LOCAL  
-OUTFMT=pco -NORX=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457 @CSN 1.1 2496 @runat 18052004.121729.20297 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLLOC=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -TERADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBdb1: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pac: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sca: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: gb\_vt: \*  
16: em\_ba: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_ov: \*  
22: em\_ox: \*  
23: em\_pac: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sy: \*  
28: em\_vt: \*

29: em\_vl: \*  
30: em\_hcg\_hum: \*  
31: em\_hcg\_inv: \*  
32: em\_hcg\_other: \*  
33: em\_hcg\_mus: \*  
34: em\_hcg\_pln: \*  
35: em\_hcg\_rtd: \*  
36: em\_hcg\_mam: \*  
37: em\_hcg\_vrt: \*  
38: em\_sy: \*  
39: em\_higo\_hum: \*  
40: em\_higo\_mus: \*  
41: em\_higo\_other: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1622	99.9	1267	8	AF452451 Glycine m
2	1453	89.5	1855	8	D26777 Citrullus v
3	1451	89.4	1247	8	AY450295 Nicotiana
4	1445	89.0	1354	8	AB029511 Solanum t
5	1435	88.4	1308	8	AF044172 Solanum t
6	1431	88.2	1281	8	SP10AL
7	1431	88.2	1303	6	E06716
8	1417	87.3	1280	8	BUY10845
9	1409	86.8	1238	8	BUY10847
10	1398	86.1	1233	8	AY063827
11	1384	85.3	1253	8	AT0ACLY
12	1381	85.1	1337	8	AK071279
13	1379	85.0	1332	8	ATCY83A
14	1378	84.9	1319	8	ZMCSOATL
15	1377	84.8	1363	8	AF073697
16	1374	84.7	966	6	AX506198
17	1374	84.7	966	6	AX651825
18	1373	84.6	1671	8	AB040503
19	1371	84.5	1355	8	BT009041
20	1366.5	84.2	1254	8	AY045825
21	1364	84.0	1371	8	WHTCYS1
22	1360	83.8	993	6	AX652865
23	1360	83.8	1280	8	AK061993
24	1360	83.8	1281	8	AK099598
25	1359	83.7	5259	6	AX653890
26	1356	83.5	1290	8	AF073695
27	1348	83.1	1368	8	AT0011976
28	1334	82.2	1192	8	AY353092
29	1337	81.1	1562	8	AK099217
30	1317	81.1	1602	8	AK065007
31	1317	81.1	2104	8	AK065652
32	1316	81.1	1378	8	AB029512
33	1314	81.0	1404	8	AF044173
34	1314	81.0	1555	8	AK104704
35	1290	79.5	953	8	CAR6024
36	1271	78.3	1377	8	AT0271727
37	1271	78.3	1460	8	AY128885
38	1271	78.3	1483	6	E08016
39	1271	78.3	1484	8	SP1CISY
40	1271	78.3	1556	8	AY099721
41	1270	78.3	1492	8	AF078693
42	1261	77.7	1210	8	AY096681
43	1261	77.7	1378	8	AT0ACLY2
44	1261	77.7	1432	8	SOCYSK
45	1261	77.7	1489	8	AY065375

RESULT 1

ALIGNMENTS

AF452451 1267 bp mRNA linear PLN 21-JAN-2002  
LOCUS Glycine max cysteine synthase mRNA, complete cds.  
DEFINITION  
ACCESSION AF452451 GI:18252505  
VERSION AF452451.1 GI:18252505  
KEYWORDS  
SOURCE  
ORGANISM  
Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabiales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
1 (bases 1 to 1267)  
Demosthenis, C. and Krishnan, H.B.  
Nucleotide sequence of a soybean cDNA encoding cysteine synthase  
Unpublished  
2 (bases 1 to 1267)  
Demosthenis, C. and Krishnan, H.B.  
Direct Submission  
Submitted (28-NOV-2001) USDA-ARS, University of Missouri, 108W  
Curtis Hall, Columbia, MO 65211, USA  
LOCATION/Qualifiers  
1. 1267  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
82..1059  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="AA166291.1"  
/db\_xref="GI:18252506"  
/translation="MAVERSGIAKDVTEIGTPIVYINKLADGCVARVAKLELMEP  
CSVVKRIGTSMADBEKLTITGKSVLEIFPSGNTGICLMAAARGKILITMPA  
SWSIKRITLILAFGAELVITDPKAKMGAAQKAEILAKTPNATILQSPENPAPVH  
YETTGPEIWRGSGDKIDAFVSGIGTGITGAGKYLKEQNPMLKIGVFPVBSPIVSG  
GKQPKRIQIGIAGPIPGVLENLDEVIDISSDEALETKMLAKKEGIFVIGISGA  
AAAAFQAKRPENAGKILIVAFPSFGRYLSVLPFSVREASMTPEP"

## CDS

Alignment Scores:  
Pred. No.: 2,366-117 Length: 1267  
Score: 1622.00 Matches: 324  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.69% Mismatches: 0  
Query Match: 99.94% Indels: 0  
DB: 8 Gaps: 0

US-09-931-457a-31 (1-325) x AF452451 (1-1267)

QY 1 MetaValaValaGluArgSerGlyIleAlaAlaAspValThrGluLeuIleGlyIleThrPro 20  
DB 82 ATGGCTGTAAAGAGTCCGGAATTCGCAAGAGTTCGGAATTCGTAATAAACCCTCA 141  
QY 21 LeuValaValaGluArgSerGlyIleAlaAlaAspValThrGluLeuIleGlyIleThrPro 40  
DB 142 TTGATATATCTAAATAACTTCGGAATTCGTAATAAACCCTCA 201  
QY 41 LeuMetGluProCysSerSerValaAspArgIleGlyIleThrSerMetIleAlaAspAla 60  
DB 202 TTGATGAGGCAATCTCTAGTGTGAAGACAGGATTCGATATGATGATTCGTAATGCA 261  
QY 61 GluGluValSerGlyLeuIleThrProGlyIleValSerValLeuIleGluProThrSerArgIle 80  
DB 262 GAAGAGGAGGAGGACTTACCACTCGGAAAGAGTCTCTCTGAGCCCAAGAGGATTCAT 321  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgIleGlyIleThrSerMetIleAlaAspAla 100  
DB 322 ACTGCGATTCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 381  
QY 101 ProAlaSerMetSerLeuGluValArgIleIleIleLeuLeuAlaPheGlyIleAlaGluVal 120  
DB 382 CTTGCTTCTATGATCTTGAAGAGGAGATCATCTTATTCATTCATTCATTCATTCATTCATTCAT 441

QY 121 LeuThrAspProAlaIleGlyIleMetIleGlyIleAlaValaGluValaGluGluIleLeuAla 140  
DB 442 TTGACAGATCCCGTCAAGGGAATGAAGGTCTGTTCCAGAGCTCGAAGAGATTCATTCGCT 501  
QY 141 LysThrProAspAlaIleThrIleLeuGluIlePheGluAsnProAlaAspProIleValHis 160  
DB 502 AAGAGCCCAATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 561  
QY 161 TyrGluThrThrGlyProGluIleThrIlePheGlyIleSerAspGlyIleAlaPheVal 180  
DB 562 TATGAACACACTGTCACAGATATGAAGAAGCTCCGATGAGGAATGATTCATTCATTCATTCAT 621  
QY 181 SerGlyIleGlyIleThrGlyIleThrIlePheGlyIleValaGlyIleValaGlyIleVala 200  
DB 622 TCTGGATATGCACTGTCGTGTCATTAACAGAGGTCTGGAATAATCTTAAAGACAGAT 681  
QY 201 ProAsnIleLeuIleGlyIleValaGluProValaGluSerProValaLeuSerGlyIleVal 220  
DB 682 CCGATATTAACCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 741  
QY 221 ProGlyProHisIleGlyIleGluIleGlyIleGlyIleGlyIlePheIleProGlyIleVala 240  
DB 742 CTTGTCACACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGAT 801  
QY 241 AsnLeuLeuAspGluValaValaGluIleSerSerAspGluAlaIleGluThrAlaIleVala 260  
DB 802 ATCTCTTGTGATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 861  
QY 261 LeuAlaLeuGlyGluIlePheValaGlyIleSerSerGlyIleAlaAlaAlaAlaAlaAla 280  
DB 862 CTTGCTTCTATGATCTTGAAGAGGCTTATTCGGAATATCTTCGAGCTGCACTCTCTGCT 921  
QY 281 PheGluIleAlaIleValaPheProGluAsnIleGlyIleValaIleValaIlePheProSer 300  
DB 922 TTTCGATTCACAAAGACAGCAAAATTCGCGAAGCTTATTCGCTTTCCTCCAC 981  
QY 301 PheGlyIleValaIleValaPheProGluAsnIleGlyIleValaIleValaIlePheProSer 320  
DB 982 TTCGGGAGAGGATCTGCTCGGCTATTCGATTCAGTACAGACGCAAGCTGAAGAC 1041  
QY 321 MetThrPheGluPro 325  
DB 1042 ATGACTTTGAGCC 1056

RESULT 2  
LOCUS CNAPOCS7 1855 bp mRNA linear PLN 28-JAN-2003  
DEFINITION Citrullus vulgaris mRNA for cysteine synthase, complete cds.  
ACCESSION D28777.1 GI:466530  
VERSION D28777.1 GI:466530  
KEYWORDS O-acetyl-L-serine (thiol)-lyase; O-acetyl-L-serine acetylase  
(adding hydrogen sulfide); O-acetylserine sulfhydrylase.  
SOURCE  
ORGANISM Citrullus lanatus (watermelon)  
Citrullus lanatus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.  
REFERENCE  
AUTHORS Noji, M., Murakoshi, I. and Saito, K.  
TITLE Molecular cloning of a cysteine synthase cDNA from Citrullus  
vulgaris (watermelon) by genetic complementation in an Escherichia  
coli Cys- auxotroph  
JOURNAL Mol. Gen. Genet. 244 (1), 57-66 (1994)  
MEDLINE 94316193  
PUBMED 8041362  
2 (bases 1 to 1855)  
REFERENCE  
AUTHORS Saito, K.  
TITLE Direct Submission  
JOURNAL Submitted (28-FEB-1994) Kazuki Saito, Chiba University, Faculty of  
Pharmaceutical Sciences, Lab. Mol. Biol. Biotech.; Yayoi-cho 1-33,  
Inage-ku, Chiba 263, Japan (E-mail:ksaito@ph.chiba-u.ac.jp,  
Tel:81-43-290-2904, Fax:81-43-290-2905)  
COMMENT On Mar 30, 1994 this sequence version replaced gi:391655.

D17363: Submitted (05-Aug-1993) to by:  
Kazuki Saito.

```

FEATURES
source      Location/Qualifiers
1. 1855     /objectname="C:\tm1\us\lanatue

```

```

/tissue type="etiolated seedlings"
564. .1541
/END number=14 3 00 0"
CDS

```

```

/productc="cysteine synthase"
/protein_id="BAA05965.1"
/db_xref="cgi:540497"
/feature_id="MAAKSTIAKDVTEIGTFLYYLRNRVSDGVAVAALTEMBE
CSYKDRIGYGMISDAENKRLTGSGEIVTEPSGNTGAPFAIAAGYLLICMP
SMSELRPTTIFAPALVILDPBAGKAGNORAEKATGAPSLAIDGPNBANKH
YATTEPRITWRSSGSDIALVSGIGTGTGYAAKTKLEONPNIKYQEPRESAIS
GEGPHKTKIGYGAQITPEVLDVNLDEIVGVSSESLTEPAKLALTEFLVGISG
AAALRIKRGEPENGGKILVAFPSFGRIYLTSLFSEVREKLENNYFEI"

```

ORIGIN

Alignment Scores:

Pred. No.:	5,63e-104	length:	185
Score:	1453.00	Matches:	283
Percent Similarity:	94.46t	Conservative:	24
Best Local Similarity:	87.08t	Mismatches:	18
Query Match:	89.53t	Indels:	0
DB:	8	Gaps:	0

US-09-931-457A-31 (1-325) X CNAPCCS7 (1-1855)

QY	1	MetIaValaGluArgSerSelyIleAlaIaIysAspValIThrGluLeuIleIGLYsthrPro	20
Db	564	ATGGCTGATGCAGAAAGCTTACAAATACCTAAAGATGTTACTGAACATAATGGGAATACACCA	623
QY	21	LeuValITyLeuAsnLysLeuAlaAspGlyCysValaIaArgValaIaAlaIysLeuGlu	40
Db	624	CTTGATATCCCTCAACCGGTGTGTGATGGCTCCGTGGCTCCGGGTACGCCAAGTTGAG	683
QY	41	LeuMetGluProCysSerSerValIysAspArgIleGlyITySerMetIleAlaAspIa	60
Db	684	ATGATGGAGCCTTGCTGCACAGTGCMAAAGATGATGTGCTATGACATATTTGCATGCA	743
QY	61	GluGluLysGlyLeuIleIThrProGlyLysSerValIleuIleGluProThrSerGlyAsn	80
Db	744	GAAGAACAGGAGTCTTATACACTCCGGGGAGAGTGTCTGATTGAACCTTACAGTGGTAT	803
QY	81	ThrGlyIleGlyLeuAlaPheMetAlaIaIaIaArgGlyITyIysLeuIleIThrMet	100
Db	804	ACTGGTATAGGCTTGCGCTTCAATTCCTGCTGCGCAAGGTTAAAGCTTATATATTGCATG	863
QY	101	ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValaGluLeuVal	120
Db	864	CCTGATCAATGAGCTCTGAAAGAAAGAACCAATCTTCAGAGCTTTCGGAGCTGCACTGGAT	923
QY	121	LeuThrAspProAlaIysGlyMetIysGlyIaIaValaGluLysAlaGluGluIleLeuAla	140
Db	924	CTCACAGATCCAGCTCCGGGAGATGAAGAGACAGCTTCMAAGGCTGMAAGATTAAGGCA	983
QY	141	LysThrProAsnAlaTyrlIleLeuGlnGlnPheGluAsnProAlaAsnProLysValaHis	160
Db	984	AAACACCTTAATCTTACATCCTCAGCAATTTGAAAAACCTGCTAACCCGGAAGATCCAC	1043
QY	161	TyrgluIThrIThrGlyProGluIleITrpLysGlySerAspGlyLysIleAspAlaPheVal	180
Db	1044	TATAGAGCACACTGGTGCAGAGATTCGAGGGGCTCAGGTGGGAAGAAATTGACCCACTTGC	1103
QY	181	SerGlyIleGlyThrGlyGlyThrIleIThrGlyAlaGlyLysITyLeuLysGlnGluAsn	200
Db	1104	TTGGTATTAAGGACTCGAGGCTACAGTGCACAGTGCAGGGAAGATATCTCAAGACAAATAT	1163

QY 201 ProksnlelyseuilegiylValGluProValGluSerProValLeuSerGlyValys 220  
 Db 1164 CCTAATATTAAGTTATATATGCGGTTGAACCTGTGAAGTGAATTCTATCTGAGGAAAG 12233  
 QY 221 ProGlyProHisIlylllegInGlyllegIlyAlaGlyPheIleProGlyValLeuGluVal 240  
 Db 1224 CCTGGCCCACTATAAAATCCAAAGGAATGGGGCGAGGCTTCAATCCCTGAGTTTGGAATCTC 12833  
 QY 241 AsnLeuLeuSpGluValAlaGlnIleSerSerSpGluAlaIleGluThrAlaIysLeu 260  
 Db 1284 AATTGCTGAGTGAAGTTATTCAAGTATCAAGTAAAGATATTGGAACCTGCAAAAGCTG 13433  
 QY 261 LeuAlaLeuysGInGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280  
 Db 1344 CTTCATTTGAAAGAAAGGCTACTACTGTTGGTATATCATCTGGCGGCGCGCGCTGCTGCA 14033  
 QY 281 PheGlnIleAlaIybaGrProGluAsnAlaGlyIysLeuIleValAlaValPheProSer 300  
 Db 1404 ATTTGGAGTTGGCAAGAGACCGAAGAAAGCTGAGAAAGCTATTGTTGGCGATTTCCAGAT 14633  
 QY 301 PheGlyGluArgIyrlLeuSerSerValIlePheGluSerValArgArgGluAlaGluSer 320  
 Db 1464 TTTCGGGAGACATACCTCTCTCAACCGTCTGCTTCAGTCTGTGAACAGAGACACGAGAT 15233  
 QY 321 MetThrPheGluPro 325  
 Db 1524 ATGGTTTTCAGGCCA 1538

LOCUS	AY450295	1247 bp	mRNA	linear	PLN 22-NOV-2003
DEFINITION	Nicotiana glumbaginifolia cysteine synthase (Csl) mRNA, partial				

[illegible]

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 1247)	Liszewska, P. and Sirko, A.	plumbagin: cDNA encoding
		Isolation of Nicotiana glauca
		O-acetylserine (thiol) lyase and serine acetyltransferase

**AUTHORS** Liszewska, P. and SIKKO, A.  
**TITLE** Direct Substitution  
**JOURNAL** Submitted (26-OCT-2003) Institute of Biochemistry and Biophysics,  
Polish Academy of Sciences, Pawinskiego 5A, Warsaw 02-106, Poland

```

/mol_type="mrna"
/db_xref="taxon:4092"
<1..1247

```

```

/EC number="2.5.1.47"
/note="NPC1: O-acetylserine (thiol) lyase"
/codon_start=1
/codon_usage="surphase"

```

[illegible]

## Alignment Scores:

Pred. No.: 4,83e-104 Length: 1247  
 Score: 1451.00 Matches: 281  
 Percent Similarity: 95.03% Conservative: 25  
 Best Local Similarity: 87.27% Mismatches: 16  
 Query Match: 89.40% Indels: 0  
 DB: 8 Gaps: 0

US-09-931-457a-31 (1-325) x AY50295 (1-1247)

```

QY 4 GluArgSerGlyIleAlaValAspValThrGluLeuIleGlyLysThrProLeuValTyr 23
    |||
DB 4 GAAAGAGATGAAATTCGCAAGATGTAAGTAATGATCGTAACACTCTTGGGTAC 63
QY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGlu 43
    |||
DB 64 CTGAAATATGTGTGATGGGTGTGCGCCGCTTCCCAAGCTGGAAGCATGGAG 123
QY 44 ProCysSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLys 63
    |||
DB 124 CCAAGCTCTAGTGTTAAGATAGATGGATGTTATGATGATTCACAGTCTGAGAGAA 183
QY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyTyr 83
    |||
DB 184 GGCTGATCAAACTCGCAGAGAGTGTCTCATTAAGACCTCAAGTGGAAACACTGGAGTA 243
QY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleThrMetProLaser 103
    |||
DB 244 GGATGGCATTTATGGCTGCTCTTAAGGCTCAAACTCATACATACATGCTTCTTCA 303
QY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123
    |||
DB 304 ATGAGCTTGGAGGAGAAATTAATCTCCGCTTGTGTGCTGATGTTGATTAACGAT 363
QY 124 ProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGlnLysLeuAlaLysThrPro 143
    |||
DB 364 CCAAGCAAAAGGAGTGAAGGTTCTTATCAAGAGCTTAAGAAATTAAGGCCAAGACACC 423
QY 144 AsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThr 163
    |||
DB 424 AACCTCCTTATTTCTTCGCAATTTGAAACCCCTCAAAACCAAGGATACATAGAGACC 483
QY 164 ThrGlyProGluIleTyrLysSerAspGlyLysIleAspAlaPheValSerGlyTyr 183
    |||
DB 484 ACTGGTCTCTAGATCTGAAAGGCTCAATGGGAAATGATGCTCTGCTCTGAAAT 543
QY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIle 203
    |||
DB 544 GGAACAGAGGACCAATACAGGTTCCAGGAGATTTAAGAGCGAAGACCCGACATA 603
QY 204 LysLeuIleGlyValAlaGluProValGluSerProValLeuSerGlyGlyLysProGlyPro 223
    |||
DB 604 AAGCTCTATGCTGTGGAAACAGTGAAGGTCATTTCTTGAGAGAAAGCCCTGCTCA 663
QY 224 HisLysIleGlnGlyIleGlyValAlaGlyPheIleProGlyValLeuGluValAsnLeu 243
    |||
DB 664 CATTAGATCCAGGAGATTTGCTGCTGTTTCAATTCCTGTTTGAAGATGGCTTAT 723
QY 244 AspGluValAlaGlnLysSerSerAspGluAlaIleGluThrAlaLysLeuAlaLeu 263
    |||
DB 724 GATCAAGTAAATTCAGATTCAAGTGAAGAACCCCTAAGAACTGCAAGCTCTGCGCTTG 783
QY 264 LysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaAlaPheGlnIle 283
    |||
DB 784 AAGGAGAGTTTCTTGTGGAGATTCATCTGCTCTCTCTGCTGCGGACATCAACT 843
QY 284 AlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlu 303
    |||
DB 844 GCTAAGGCGCCCTGAAATGCTGGAGAGCTCATGTTCTTCTTCCCAACTCTCGGAG 903
QY 304 ArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPhe 323
    |||
  
```

```

DB 904 CCAATCTTCTCTCTGTCTCTTTGAATCTGTAGACGGGAGACAGAACATACCGTG 963
QY 324 GluPro 325
    |||
DB 964 GAGCCT 969
QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
    |||
DB 85 ATGGCGGGGAAAGATTTGGAATTCACAGAGATGTAAGTATGATGTAACACTCT 144
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
    |||
  
```

RESULT 4  
 AB029511  
 LOCUS  
 DEFINITION  
 Solanum tuberosum PCS-1 mRNA for cytosolic cysteine synthase,  
 complete cds.  
 AB029511.2 GI:32997083  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Solanum tuberosum (potato)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1  
 Maryama, A. and Ishizawa, K.  
 Cytosolic Cysteine Synthase from Potato Sprout  
 Published Only in Database (2001)  
 2 (bases 1 to 1354)  
 Maryama, A. and Ishizawa, K.  
 Direct Submission  
 Submitted (02-JUL-1999) Akiko Maryama, Chiba University, Faculty  
 of Pharmaceutical Sciences, Laboratory of Molecular Biology and  
 Biotechnology, Research Center of Medicinal Resources; Yayoi-cho  
 1-33, Inage-ku, Chiba, Chiba 263-8522, Japan  
 E-mail: amaryama@chiba-u.ac.jp, Tel: 81-43-290-2906,  
 Fax: 81-43-290-2905  
 On Jul 19, 2003 this sequence version replaced gi:12081916.  
 COMMENT  
 FEATURES  
 source  
 1..1354  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /strain="Dansyaku"  
 /db\_xref="taxon:4113"  
 /cell\_type="sprout"  
 /note="cetruploid"  
 1..1354  
 /gene="PCS-1"  
 85..1062  
 /gene="PCS-1"  
 /product="cytosolic cysteine synthase"  
 /codon\_start=1  
 /protein\_id="BAB20861.1"  
 /db\_xref="GI:12081917"  
 /translation="MAGEKIGIAKDVTETLIGNTPLVYINNVVVGCVARVARELSEMP  
 CSYVDKRGISMTDPAERKGLKPGESVLIETSGNTGVLAFMAAGAYKLIITMPS  
 SMSLRRITIRAFKALVLTDPAKMKKASAEIKAKIPNSYILQOFENPANKIH  
 YETGPEITKSGNAGVDAIVSGITGGITTSKGLRQNPVVKLYGVEPVSSATLSG  
 GPGPHKIOGIGAPTPGVLEYNLIDDPVYVSSSISIKALIAKEGLIVSSGAA  
 AAALKVAPRPAKGLIVLIPSPFERLSSLVFETVREKENVTEP"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.57e-103 Length: 1354  
 Score: 1445.00 Matches: 279  
 Percent Similarity: 93.85% Conservative: 26  
 Best Local Similarity: 85.85% Mismatches: 20  
 Query Match: 89.03% Indels: 0  
 DB: 8 Gaps: 0  
 US-09-931-457a-31 (1-325) x AB029511 (1-1354)

Db 145 TTGGTATACCTGAATAATTTGGTGGATGGGTGTGTGCAAGTGTGCTGCAGGCTGGAA 204  
 Qy 41 LeuMetGluProCysSerSerVallyAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 205 AACATGGAGCCATGCTCTAGTGTGAAGATGAGATCGGTATAGTATGATTCAGATGCT 264  
 Qy 61 GluGluValyGlyLeuIleThrProGlyLysSerValIleuIleGluProThrSerGlyAsn 80  
 Db 265 GAGGAGAGAGGCTCTATCAACCCGCGAGAGTCTCTCATGCACTTCAGAGTGAAC 324  
 Qy 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
 Db 325 ACTGGTGTAGATTCGATTCATGCTGCTGCTGAAGGCTACAACTCATTCATTCAGAGT 384  
 Qy 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuMetAlaPheGlyValGluLeuVal 120  
 Db 385 CCTCTTCATATGAGCTCTGAGAGAGATATTCCTCGTCTTCGCTGTGAGTGTGCTG 444  
 Qy 121 LeuThrAspProAlaLysGlyMetLysGlyAlaAlaValGlnLysAlaGluGluIleLeuAla 140  
 Db 445 CTACCGATCCAGCAAAAGGAGTGAAGGGTGTATTCAAAGCTGAGAGGATTAAGGCC 504  
 Qy 141 LysThrProAlaAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
 Db 505 AAAATACCCCACTCTATATCTTCAGCAATTTGAAACCTGCTAACCCCAAGATACAC 564  
 Qy 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
 Db 565 TATGATACCACTGCTCTGAGATCTCGAAAGCTCAATGCGAAGATGATCTCTAGTC 624  
 Qy 181 SerGlyIleGlyThrGlyLysThrIleThrGlyValGlyLysThrLeuLysGluIleAsn 200  
 Db 625 TCTGGATTTGGAACAGAGGACAGATACGTTCTGAGGCAAGTATTTGAGAGGCAAGAC 684  
 Qy 201 ProAlaIleLysLeuIleGlyValGluProValGlnSerProValLeuSerGlyLys 220  
 Db 685 CCCAAGCTTAAAGCTGTATGCGGTGAGAACCAAGTTGAAAGTCTATCTTCTGTGGAAG 744  
 Qy 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 Db 745 CCGGTCCACATTAAGATTCAGGGGATTCGGTGTGCTGCTCTGCTGTGTAAGT 804  
 Qy 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 Db 805 AACCTTATGATATGATGATGATTCAGGTTTCAAGTGAAGATCCATTAAGATGCTTAAGCT 864  
 Qy 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaAla 280  
 Db 865 CTGGCATTTGAAGAGATTTGCTAGTGGAAATTTCAATCTGCTGCTGCGGGCGCGGCA 924  
 Qy 281 PheGluIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
 Db 925 ATTAAGATCCCTTAAGGCGCTGAAATGCTGGAGAACTCATTTGTTTATTTCCCAAC 984  
 Qy 301 PheGlyLysArgTyrLeuSerSerValIlePheGluSerValAlaArgArgIleAlaGluSer 320  
 Db 985 TTGCGAGAGGAGATCTCTTCTCTGTAATCTTCGAAACTCTCAGAGAGAGAGAGAAC 1044  
 Qy 321 MetThrPheGluPro 325  
 Db 1045 ATGACTGTGAGCCT 1059  
 RESULT 5  
 AF044172 1308 bp mRNA linear PLN 06-JUL-1998  
 LOCUS Solanum tuberosum cysteine synthase mRNA, complete cds.  
 DEFINITION  
 AF044172 GI:3290019  
 ACCESSION  
 AF044172.1  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,  
 Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots;

asterides; lamids; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 1308)  
 Heese, H. and Hoeftgen, R.  
 Isolation of cDNAs encoding cytosolic (Accession No. AF044172) and  
 plastidic (Accession No. AF044173) cysteine synthase isoforms from  
 Solanum tuberosum (RGR98-057)  
 Plant Physiol. 116, 1604 (1998)  
 Heese, H.  
 2 (bases 1 to 1308)  
 Direct Submission  
 Submitted (22-JAN-1998) Freie Universitaet Berlin, Institut fuer  
 Angewandte Genetik, Albrecht-Thaer-Weg 6, Berlin 14195, Germany  
 Location/Qualifiers  
 1..1308  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cuffilter="Bertolina"  
 /db\_xref="taxon:4113"  
 35..1012  
 /note="CS-A; O-acetylserine (thiol) lyase; cytosolic  
 isoform"  
 /codon\_start=1  
 /product="cysteine synthase"  
 /protein\_id="PAC25635.1"  
 /db\_xref="GI:3290020"  
 /translation="MAGKRIQIAQDVTEIGNTPLVLTNVVDGCVARVAKLSEMP  
 CSVDRIQYSMTDABEKQIKPESVLEPTSGNTVGAFMAAAGVYKILTPS  
 SMLEKRIIRGSEELVTPAKMKALIKAEKIKAKTPSYLLQDFENPKIH  
 YETTGPEIKWSNGVDALASIGGTYGSGKTLRQNEPVKLYGVEPVESALLSG  
 KRGPBKIGIGAGPIPGVLEVNLIIDVVQVSSSISMAKILAKGLVIGISGAA  
 AAAAIKAKRPRNAKLIIVTFPSRGRHYSVLEFETVRRRAENTVP"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,056-103 Length: 1308  
 Score: 1435.00 Matches: 277  
 Percent Similarity: 93.23% Conservative: 26  
 Best Local Similarity: 85.23% Mismatches: 22  
 Query Match: 86.42% Indels: 0  
 DB: Gaps: 0  
 US-09-931-457a-31 (1-325) x AF044172 (1-1308)  
 Qy 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20  
 Db 35 ATGGCGGGGGAAGAAGATTGGAATTCGCAAGATGTAAGTAAATTTGTTAACTCTCT 94  
 Qy 21 LeuValIyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
 Db 95 TTGGTATACCTGAATAATGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 154  
 Qy 41 LeuMetGluProCysSerSerVallyAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 155 AACATGGAGCCATGCTCTAGTGTGAAGATGAGATCGGTATAGTATGATTCAGATGCT 214  
 Qy 61 GluGluValyGlyLeuIleThrProGlyLysSerValIleuIleGluProThrSerGlyAsn 80  
 Db 215 GAGGAGAGAGGCTCTATCAACCCGCGAGAGTCTCTCATGCACTTCAGAGTGAAC 274  
 Qy 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
 Db 275 ACCGCTGTAGATTCGATTCATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 334  
 Qy 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuMetAlaPheGlyValAlaGluLeuVal 120  
 Db 335 CCTCTTCATATGAGCTCTGAGAGAGATATTCCTGCGTGTGCTGCTGCTGCTGCTGCTG 394  
 Qy 121 LeuThrAspProAlaLysGlyMetLysGlyAlaAlaValGlnLysAlaGluGluIleLeuAla 140  
 Db 395 CTACCGATCCAGCAAAAGGAGTGAAGGGTGTATTCAAAGGCTGAGAGATTAAGGCC 454  
 Qy 141 LysThrProAlaAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160

D	b		455	AAAACACCACCTCCTATATTCTTCAGCAATTGGAAACCCCGTCAACCCAAAGATCAC	514
Oy			161	TyrGluThrThrglyProGlnIleTrpIysGlySerAspGlyValLeaspAlaPheVal	180
D	b		515	TATAGAACCATCGTGCTCTGAGATCTGGAAAGGCTCAAAACGGGAAATGGAGACTTAGGCC	574
Oy			181	SerGlyIleGlyThrGlyValThrIleThrGlyValagIlysTyrlsLeuValGIslasn	200
D	b		575	TTGGATTGGAAACGAGGAGCAACGATACCTGGTTCAGCCAAGTATTTGAGAGACAGAAC	634
Oy			201	ProAsnIleLysLeuIleGlyValGlnProValGlnSerProValLeuSerGlyValYls	220
D	b		635	CCCAACGTAAAGCTGATAGCGCTGAACACAGTGAAGAATGCATACCTTCGTGGTGAAG	694
Oy			221	ProGlyProHisValIleGlnGlyValIleValagIylsPheIleProGlyValLeuGlnVal	240
D	b		695	CCTGGTCCACATAAATTCACAGGGGATGGTGTGATTTCATTCCTGGAGTTTGGAAGTT	754
Oy			241	AsnLeuLeuAspGlnValValGlnIleSerSerAspGlnAlallegluThrAlaLysIleu	260
D	b		755	AACCTTATATGATGATGTAATGTTCAAGTTCAAGTGAAGATCATAGAAATGGCTTAGGTT	814
Oy			261	LeuAlaLeuLysGlnGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaAla	280
D	b		815	CTGGCATTTGAAGAGAGGATGCTGATGATGGAAATTCATCTGGTGTCTGGCCGGCGGCA	874
Oy			281	PheGlnIleAlaLysArgProGlnAsnIleagIylsLeuIleValAlaPheProSer	300
D	b		875	ATTAAAGTGTCTAACGCCCTGGAATCTGGAAAGCTCATTTGTTGTAATTTCCCAAGC	934
Oy			301	PheGlyGlnArgTyrLeuSerSerValIleuPheGlnSerValArgArgGlnAlaGlnSer	320
D	b		935	TTGGAGAGGACATATCTTCTCTGTTACTCTTCGAAACTGTGCACAGAGAGCAGAGAAC	994
Oy			321	MethrPheGlnPro	325
D	b		995	ATGATGTGGAGCCT	1009
<b>RESULT 6</b>					
LOCUS	SPOL.				
DEFINITION	Spinach mRNA for O-acetylserine(thiol)-lyase, complete cds.				
ACCESSION	D10476				
VERSION	D10476.1 GI:218276				
KEYWORDS	O-acetyl-L-serine acetate-lyase (adding hydrogen sulfide); O-acetylserine(thiol) lyase; cysteine synthase. Spinacia oleracea ('spinach')				
SOURCE	Spinacia oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Spinacia.				
ORGANISM	1 (bases 1 to 1281) Saito, K., Miura, N., Yamazaki, M., Hizano, H. and Murakoshi, I. Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8078-8082 (1992)				
JOURNAL	Medline				
PubMed	92390392				
REFERENCE	2 (bases 1 to 1281) Saito, K.				
AUTHORS	Direct Submission Submitted (12-FEB-1992) Kazuki Saito, Chiba University, Faculty of Pharmaceutical Sciences, Lab. Mol. Biol. Biotech.; Yanoi-cho 1-33, Inage-ku, Chiba, Chiba 263, Japan (E-mail:ktsaitoe@chiba-u.ac.jp, Tel:045-290-2904, Fax:043-290-2905)				
TITLE	Submitted (12-FEB-1992) to DDBJ by: Faculty of Pharmaceutical Sciences, Chiba University Yanoi-cho 1-33 Chiba 260 Japan Phone: 0472-51-1111 x2707 Fax: 0472-55-1574.				Kazuki Saito

FEATURES	Source	Location/Qualifiers
		1..1281
		/organism="Spinacia oleracea"
		/mol_type="mRNA"
		/db_xref="caxon:3562"
CDS		52..1029
		/EC_number="4.2.99.8"
		/codon_start=1
		/product="O-acetylserine (thiol) lyase"
		/protein_id="BA01279.1"
		/db_xref="gi:218277"
		/translation="MVEKATLADTVELGKTPLYLVLTVAQCARVAKKEGHE SSVDRIRGFPMITDAESGLITPESVLETPSGNNGIGLATIAAKGKLTITP CSVLRRITLRAFGAHLILTPAGMKGAQVAKARIKIDPNSTIIQQFNPANPKV YETPGEIWKQXGKIDITFSGIGTGGITITGAGKYLEKQNDVLTIGLPEVASVLS YTPGPHKIQGGAGNITPGVILVNTIIDEVVOISSESEIMKALALKEGVLGSGA AAALIKVAKRPMAGKILVAVFPPSGERYLSVLPDVRKASMTWES"
		1281
polya_site		
ORIGIN		
Alignment Scores:		
Pred. No.:	1,81e-102	Length: 1281
Score:	1431.00	Matches: 281
Percent Similarity:	94.14%	Conservative: 24
Best Local Similarity:	86.73%	Mismatches: 19
Query Match:	88.17%	Indels: 0
DB:	8	Gaps: 0
US-09-931-457A-J1 (1-325) x SPIOLK (1-1281)		
OY	1 MetAlaValAlGluArgSerGlyLLeuAlaYAspAlaThrGluLeuIleGlyLysThrPro	20
DB	52 ATGGTTGAGGAGAGAGGCGCTTCCTTGCCTAAAGTGTGACTGAATGATGGGAAAGCGCA	111
OY	21 LeuValTYrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaYsLeuGlu	40
DB	112 TTGGTATATCTCAACACTGCTCCCAATGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGAA	171
OY	41 LeuMetGluProCysSerSerValYAspArgIleGlyTYrSerMetIleAlaAspAla	60
DB	172 GGAATGGAACCTTGCTCTGATGTATAAGACAGGATGGTGGTCACTAAGATATCTGAAGCT	231
OY	61 GluGluLysGlyLeuLeuLeuThrProGlyLysSerValLeuIleGluProThrSerGlyAsn	80
DB	232 GAAAGAAAGGAGGCGCTTATTCACCTCGAAGAGGTCTCTGATTGAGCCACCAAGCGAAT	291
OY	81 ThrGlyTYrLeuAlaPheMetAlaAlaAlaArgGlyTYrYsLeuIleIleThrMet	100
DB	292 ACTGGCATTTGATTACCTTCATCCGACAGCTAAAGGTTCAAGCTCATCATTCATCGATG	351
OY	101 ProAlaSerMetSerLeuGluArgArgIleLeuLeuAlaPheGlyValAlaGluLeuVal	120
DB	352 CCGACATCATGATGTTTGAAGCGAGAACTATCTCAAGGCGCTTTGGTGTGAGCTTATC	411
OY	121 LeuThrAspProAlaYsGlyMetLysGlyAlaValAlaGluYsAlaGluGluIleLeuAla	140
DB	412 CTTACTGATCCAGCAAAAGATGAAAGGCGCTGTTCAAGAGCTGAGAGATCCGTGAC	471
OY	141 LysThrProAsnAlaTYrIleLeuGluGluIlePheGluAsnProAlaAsnProYsValHis	160
DB	472 AAAACCTCCAAATTCATATATCTCAACAAAGTTGAAAACTCGCAACCCCAAGGTTCAT	531
OY	161 TYrGlyThrThrGlyProGluIleTrrPylsGlySerAspGlyYsIleAspAlaPheVal	180
DB	532 TTGGAACAACCTGACACGAATTTGGAAAGGCAAGGTGAAAAATTGATATATTCGTC	591
OY	181 SerGlyTYrIleGlyThrGlyGlyThrIleThrGlyAlaGlyLYrLeuYsGluGluAsn	200
DB	592 TCTGGATATGGGACTGGAGATACAAATTAACAGGTGCAAGAAATATCTAAAGGAACAAAC	651
OY	201 ProAsnIleLysLeuIleGlyValAlaGluProValGluSerProValLeuSerGlyLYs	220
DB	652 CCGAATGTTAAGCTATTTGCTCTGAAACCAATGCAAAATGCTGATATTTGTCGTGAGGAAAA	711

[illegible]

RESULT 7				
E06716				
LOCUS	E06716	1303 bp	RNA	linear
DEFINITION	cDNA encoding Spinacia cytochrome synthetase.			
ACCESSION	E06716			
VERSION	E06716.1	GI:2174898		
KEYWORDS	JP 1994038770-A/1.			
SOURCE	Spinacia oleracea (spinach)			
REFERENCE	PAT 29-SEP-1997			

ORGANISM	REFERENCE
Spinacia oleracea	
Eubryophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
Caryophyllales; Amaranthaceae; Spinacia.	
1 (bases 1 to 1303)	
Saitou,K. and Murakoshi,I.	
GENE CODING CYSTEINE SYNTHETASE	
PATENT: JP 1994038770-A 1 15-FEB-1994;	
MITSUBISHI CORP, MITSUBISHI KASEI CORP	
OS Spinacia oleracea	

PN	JP 1994038770-A/1
PD	15-FEB-1994
PF	05-FEB-1992 JP 1992020315
PI	SAITOU KAZUSUE, MURAKOSHI ISAMU
PC	C12N15/60,C12N9/88,(C12N9/88,C12R1:19);
CC	strandedness: Double;
CC	topology: linear;
CC	hypotheical: No;
CC	anti-sense: No;
CC	*source: strain=parade;
CC	*source: clone=PTV118N;
FH	key
FH	Location/Qualifiers
FT	5'UTR
FT	1..51
FT	52..1029
FT	/product='Spinacia cysteine synthetase' FT
3'UTR	1030..1303.

```

FEATURES
  source
    location/Qualifiers
    1..1303
    /organism="Spinacia oleracea"
    /mol_type="genomic RNA"
    /db_xref="taxon:3562"

ORIGIN

Alignment Scores:
Pred. No.:      1,85e-102      Length:      1303
Score:          1431.00        Matches:      261
Percent Similarity: 94.14%      Conservative: 24
Best Local Similarity: 86.73%    Mismatches:  19
Query Match:     88.17%         Indels:       0

```

DB:	6	Gaps:	0
US-09-931-457A-31 (1-325) x E06716 (1-1303)			
QY	1	MetalValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro	20
Db	52	ATGGTATGAGGAAAGGCCCTTCATTCGCTAAAGATGTGACTGAATTGATGGGAAAAACGCCA	111
QY	21	LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu	40
Db	112	TTGGTATATCTCAACACTGTGCGCGAGGTGTGTCTCTCGTGTTCGCGCAAAGCTGGA	171
QY	41	LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla	60
Db	172	GGATGTGAACCTTCTCTAAGTTAAAGACACGAGTTGGGTTCAGATATGATTACGATCT	231
QY	61	GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn	80
Db	232	GAAATTAAGCGGCTTATTACCTCGAGAGAGGTCTCGATTAGACCCACAGTCGGAAT	291
QY	81	ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet	100
Db	292	ACTGGCATTTGATTAAGCTTCATCGACGACAGCTAAAGGTATCAAGCTCATATACGATG	351
QY	101	ProLysMetSerArgGluThrThrTrpIleIleLeuLeuAlaPheGlyValGlyLeuVal	120

[illegible]

BUY10845 1280 bp mRNA linear PLN 02-FEB-1999  
 LOCUS BUY10845  
 DEFINITION Brassica juncea mRNA for O-acetylserine(thiol) lyase, clone OAS-TL4.  
 ACCESSION Y10845  
 VERSION Y10845.1 GI:2243119  
 KEYWORDS O-acetylserine(thiol) lyase.  
 SOURCE Brassica juncea  
 ORGANISM Brassica juncea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1  
 AUTHORS Schaefer H.J., Haag-Korner, A. and Rauech, T.  
 TITLE CDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy-metal accumulator Brassica juncea L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
 JOURNAL Plant Mol. Biol. 37 (1), 87-97 (1998)  
 MEDLINE 98281577  
 PubMed 9820267  
 REFERENCE 2 (bases 1 to 1280)  
 AUTHORS Schaefer, H.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg, Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg, FRG

FEATURES  
 source location/Qualifiers  
 1..1280  
 /organism="Brassica juncea"  
 /mol\_type="mRNA"  
 /sub\_species="Vilcasso"  
 /db\_xref="taxon:3707"  
 /clone="OAS-TL4"  
 /tissue\_type="root"  
 /clone\_lib="Lambda ZAP Express"  
 /dev\_stage="mature"  
 74..1042  
 /codon\_start=1  
 /product="O-acetylserine(thiol) lyase"  
 /protein\_id="CAA71798.1"  
 /db\_xref="GI:2243120"  
 /db\_xref="GenBank:U023733"  
 /db\_xref="GenBank:U023733"  
 /translation="MARRKADVTBLIGTPLYLNNAEGCVGVAAALNEMEPCCS VMDRIGFMSIDAEQKGLKPKESVLIETSGNTGVGLAFTAAKGVKLIITPMASMS VERRILLAFGLVLDPAKCKGKAIARAELIATPNVGYLQEPENAFRIHYET TGEPIKWDGKIDGFSVSGIGGTTTGAGKYLEKONPVKLYGVPEPISAILSGKPGPHKIGIGAGRPISYLEVDLIDREVVSDESIDMARLLAKELVGISSGAAAAA ATLARCPENAGCLPVAIVPSPGRRLISTVLDPATKKEASMTFGA"

CDS  
 222..101  
 1417..00  
 Percent Similarity: 94.36%  
 Best Local Similarity: 84.95%  
 Query Match: 87.31%  
 DB: 8  
 Gaps: 0

Alignment Scores:  
 Pred. No.: 2,22e-101 Length: 1280  
 Score: 1417.00 Matches: 271  
 Conservative: 30  
 Mismatches: 18  
 Indels: 0  
 Gaps: 0

US-09-931-457A-31 (1-325) x BUY10845 (1-1280)

QY 6 SerGlyTlleAlaLysAspValThrGluLeuIleGlyLeuThrProLeuValTyrLeuAsn 25  
 DB 80 TCTCGAATTCCTAAAGATGTTACCGAATTTATGGGAACCTCATGGTGTTCTGAAAC 139  
 DB 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuThrGluProCys 45  
 QY 140 AATGGCTGAAGAGATGTTGGCGGTCGCTGCTAAGCTCGAGATGATGAACTTGC 199  
 QY 46 SerSerValLysAspArgGlyIleGlyTyrSerMetIleAlaAspAlaGluGluValGlyLeu 65  
 DB 200 TCCAGTGTCAAGACAGAGATGGTTTATGCAATCTCTGATGCAAGCAAAAGGCTCTG 259

QY 66 TleThrProGlyLysSerValLeuIleGluProThrSerGlyValThrGlyLeu 85  
 DB 260 ATCAACAACAGAGAGAGATGCTGATTAAGCCAAAGATGGAAACAGCGAGTTGGGTTA 319  
 QY 86 AlaphMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105  
 DB 320 GCGTTACAGCGCGCGCTCCAGAGGGGTRCAAGCTCATTAATCATGCAAGCTTCAATGAGT 379  
 QY 106 LeuGluThrGlyIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125  
 DB 380 GTGGAGAGAGATCATCTCTTACCTTTTGAGTTGAGTTGCTTCAAGACCTTCTCT 439  
 QY 126 LysGlyMetLysGlyValAlaGluLeuAlaGluGluIleLeuAlaLysThrProAla 145  
 DB 440 AAGGGATGATTAAGCGCTATTCACCAAGCCCAAGAGATTTAGCCAAACCCCATGTGT 499  
 QY 146 TyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGly 165  
 DB 500 TACATGCTCCAAACAGTTTGAAGAACCTGCCAACCAAGATACACTATGAGCTACTGA 559  
 QY 166 ProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThr 185  
 DB 560 CCGTAGATATGGAAGGAGCTGATGAGAAATTCATGATGATTTGTTCTGGGATTTGACT 619  
 QY 186 GlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAlaLysLeu 205  
 DB 620 GGTGTATCAATCAACAGTGAAGAGAGATCTCAAGAAACCAAGACCTTACGTTAAGCTG 679  
 QY 206 IleGlyValGluProValGluSerProValLeuSerGlyLysProGlyProHisLys 225  
 DB 680 TATGAGAGAGAGAGAGAGATGAAAGTCTTATTCCTGGTGGAGAGAGAGAGAGAGAG 739  
 QY 226 IleGlnGlyIleGlyValAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGlu 245  
 DB 740 ATTCAGAGGATTTGGTGGCTTATACAGATGTTGAGAGGTTGATCTTATATATGAA 799  
 QY 246 ValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuLeuValGlu 265  
 DB 800 GTTGTTCAGATTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859  
 QY 266 GlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAlaAlaPheGlnIleLys 285  
 DB 860 GGTCTTCTTGGGAAATATCATCTGTCACAGCGCGCGAGCATCAATCAATCTTCAAG 919  
 QY 286 ArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluThrGly 305  
 DB 920 AGCCAGAGAAATGCTGGAGAGCTTTTGTGGCGGTTCCTGAGTTGGAGAGAGGTAC 979  
 QY 306 LeuSerSerValLeuPheGluSerValAlaArgArgGluAlaGluSerMetThrPheGlu 324  
 DB 980 CTGTGCACTGATCTTGTGATGCCAAGAGAAAGAGAGATCATATACCTTCCAG 1036

RESULT 9  
 BUY10847 1238 bp mRNA linear PLN 01-FEB-1999  
 LOCUS BUY10847  
 DEFINITION Brassica juncea mRNA for O-acetylserine(thiol) lyase, clone OAS-TL6.  
 ACCESSION Y10847.1 GI:2243123  
 VERSION Y10847.1  
 KEYWORDS O-acetylserine(thiol) lyase.  
 SOURCE Brassica juncea  
 ORGANISM Brassica juncea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1  
 AUTHORS Schaefer, H.J., Haag-Korner, A. and Rauech, T.  
 TITLE CDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy-metal accumulator Brassica juncea L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
 JOURNAL Plant Mol. Biol. 37 (1), 87-97 (1998)



MEDLINE 98281577  
 PUBMED 9620267  
 REFERENCE 2 (bases 1 to 1238)  
 AUTHORS Schaefer, H.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg,  
 Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg,  
 FRG

FEATURES  
 source Location/Qualifiers  
 1..1238  
 /organism="Brassica juncea"  
 /mol\_type="mRNA"  
 /sub\_species="vittasso"  
 /db\_xref="taxon:3707"  
 /clone="OAS-TL6"  
 /issue\_type="root"  
 /clone\_lib="lambda ZAP Express"  
 /dev\_stage="mature"  
 18..992  
 /codon\_start=1  
 /product="O-acetylserine(thiol) lyase"  
 /protein\_id="CAA71800.1"  
 /db\_xref="GI:2243124"  
 /db\_xref="SWISS-PROT:O23735"  
 /translation="MASSRIGANDVTEIIGNTPLVYNSVAEGCVRAVAKLEMEPC  
 SSVKRIQSPMSIDKKGLIKPGEVLIPTSGNTGGLAFTAAKGYKILTMPAS  
 MSIRERILLARVELVITDPAKCKKKAIAERILATPGLVLAQEPNPKIHY  
 ETTGPEIKTKTGKIDGFIISGICGTGTGAKTIKQNPVYKYGEPVSATISGS  
 KPGPKIGIGAGLIPNVEITLIDVIVQSSDIDAKLLABSGLVGISGAIA  
 AALRIARPERNAGLFAVAPVPSFGRYLSTVLPATRKESMTPEA"

## ORIGIN

## Alignment Scores:

Pred. No.: 9,936-101 Length: 1238  
 Score: 1409.00 Matches: 269  
 Percent Similarity: 93.44% Conservative: 30  
 Best Local Similarity: 84.06% Mismatches: 21  
 Query Match: 86.81% Indels: 0  
 Gaps: 0

US-09-931-457a-31 (1-325) x B0Y10847 (1-1238)

QY 5 ArgSerGlyIleAlaIlyAspValThrGluLeuIleGlyLysThrProLeuValTyrLeu 24  
 DB 27 CGTATGTGATTTGCTAATGATGATGCGAATTTGATGGAAACCTCGTGGTATCTG 86  
 QY 25 AsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluPro 44  
 DB 87 AACAGTGTTCGTAAGAGATGTTGGTGGCGTGGCGGTACGTGATGATGAGAACCC 146  
 QY 45 CysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGly 64  
 DB 147 TGCTCCACCGCTCAAGACAGAGATGGTTTACATGATGATCTGATGACGAGAAAGGGGT 206  
 QY 65 LeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGly 84  
 DB 207 CTGATTTAAACCGAGAGAGATGTGTTGATTTAGCCCAAGAGATGAGAAACCGAGATTGG 266  
 QY 85 LeuAlaPheMetAlaAlaAlaArgIlyTyrIleLeuIleIleThrMetProAlaSerMet 104  
 DB 267 TTGGCATTCACGGCGGTGCGCAAGGGCTTATATATATATATATATATATATATATAT 326  
 QY 105 SerLeuGluArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspPro 124  
 DB 327 AGTATGTGAGAGAGATATCTTCTTACCTTGTGTTGTTGTTGTTGTTTAACTGACCG 386  
 QY 125 AlaLysGlyMetLysGlyAlaValAlaGluLysAlaGluGluIleLeuAlaLysThrProAsn 144  
 DB 387 GCCAAGGCGCAAGAAAGCGCTATCCCAAGGCGGAGAGATTTGGCCAAACCCCAAT 446  
 QY 145 AlaTyrIleLeuGluIlePheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164

DB 447 GCTTACATGCTCCACAGCTTTGAGAACCTGTGTAACCAAGATCCATATGAGACTACT 506  
 QY 165 GlyProGluIleTyrPlyGlySerAspGlyLysIleAspAlaPheValSerGlyIleGly 184  
 DB 507 GACCTGATGATTTGAAAGGACAGAGAGTAAATCGATGCTTATATTTCTGGAGATTGGT 566  
 QY 185 ThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAsnProAsnIleLys 204  
 DB 567 ACTGTGTACCATACAGAGTCTGGAGAGATCTTCAAGAACAGAACCCCAAGCTTAAAG 626  
 QY 205 LeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyProHis 224  
 DB 627 CTGATGTGTGAGAGCAGTCAAGCGCTATCTTATCTGATGTAAGCCGAGCCCTTAC 686  
 QY 225 LysIleGluGlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAsp 244  
 DB 687 AAGATCAAGGAGATAGTGGCTGTGTTTATACCAATGATTTGAGACTATATCTTATGAT 746  
 QY 245 GluValValGluIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLys 264  
 DB 747 GAAGTTGTCAAGTTCAAGTGAATGATGATTCATGACATGCGAAGCTTCTGTCGCGAA 806  
 QY 265 GluGlyLeuPheValGlyLysSerSerGlyValAlaAlaAlaAlaAlaPheGluIleAla 284  
 DB 807 GAAGTCTTCTTGTGGAGATATATCCGGTGCAGACAGCTGACAGCAATCAACTTGA 866  
 QY 285 LysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyLysArg 304  
 DB 867 AAGAGCGCAAGAAACGCGGAGAACCTTTCGTGCAAGTGTCCCAAGTTTCGAGAGAG 926  
 QY 305 TyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
 DB 927 TACCTGTGACCGTACTGTTCATGATGACCAAGAGAAAGACAGATCATGATCCCTTGA 986

## RESULT 10

AY063827 LOCUS 1233 bp mRNA linear PLN 23-SEP-2002

Arabidopsis thaliana At3g22460 mRNA sequence.

AY063827 ACCESSION AY063827.1 GI:17380705

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS





**TITLE** Satto, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
**JOURNAL** Science 301 (5631), 376-379 (2003)  
**MEDLINE** 22752973  
**PMID** 12869764  
**REFERENCE** 2 (bases 1 to 1337)  
**AUTHORS** Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Doi, K.,

QY 321 MetThrphegJupro 325  
DB 1074 ATGCTGTGAGACCA 1088

RESULT 13

ATCY3A  
LOCUS ATCY3A 1332 bp mRNA linear PLN 23-AUG-1995  
DEFINITION A.thaliana mRNA for cysteine synthase.  
ACCESSION X84097  
VERSION X84097.1 GI:804949  
KEYWORDS cys-3A gene; cysteine synthase.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
1 Barroso, C., Vega, J.M. and Gotor, C.  
A new member of the cytosolic O-acetylserine(thiol) lyase gene family in Arabidopsis thaliana  
FEBS Lett. 363 (1-2), 1-5 (1995)  
MEDLINE 95246850  
PUBMED 7729527  
REFERENCE  
2 (bases 1 to 1332)  
Gotor, C.  
Title Submission  
Submitted (25-JAN-1995) C. Gotor, Inst. de Bioquímica vegetal y Fitoquímica, Dept. de Química, Univ. de Sevilla, A.C.553, R0 41080 SEVILLA, SPAIN  
Location/Qualifiers  
1..1332  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
1..1332  
/gene="cys-3A"  
99..1067  
/gene="cys-3A"  
/BC\_number="4.2.99.8"  
/codon\_start=1  
/product="cysteine synthase"  
/protein\_id="CA58893.1"  
/db\_xref="GI:804950"  
/db\_xref="GOA:P47998"  
/db\_xref="SWISS-PROT:P47998"  
/translation="MASTRADYTELIGTPIVLYNVANVAGCVGVAALKLEMEPCSS VDRIGFPMISPAEKGLIKPESVLIPTSGNTGVGLAFTPAAGCYTLITTPASMS TERRIILLAGVELVLTDPKAKGALAKAEITLAKTPRGVYLQFENPANPKIHYET TGPELWKTGKGIDFVSGITGGTITGAKRTLKQNNVXKYGVEPVESALISGKPGPHKOGICAGPTPSLVNVDLIDEVVOVSDSIDMAKQALAKELVIGISSGAEAAA AKTAAQRPENAKLFLVAFPPSPERYLSTVLPDARKAEAMTFERA"  
234..236  
/gene="cys-3A"  
/bound\_moiety="pyridoxal"

ORIGIN  
misc\_binding

Alignment Scores:  
Pred. No.: 2.13e-98 Length: 1332  
Score: 1379.00 Matches: 265  
Percent Similarity: 92.48% Conservative: 30  
Best Local Similarity: 83.07% Mismatches: 24  
Query Match: 84.97% Indels: 0  
Gaps: 0

US-09-931-457a-31 (1-325) x ATCY3A (1-1332)

QY 6 SerGlyIleAlaIysAspValThrGluLeuIleGlyIleThrProLeuValTyrLeuLeu 25  
DB 105 TCGAGATTCGCTTAAGATGTAAGTGAATGGGAACTCCATTCGTGTGATTTGAAAC 164

QY 26 LysLeuAlaAspGlyCysValAlaArgValAlaIleAlaIleLeuGluLeuMetGluProCys 45  
DB 165 AATGTTGCTGAAGAGATGTTGCTGTGTGCTGAAGCTTGAGATGATGAGAACCGTGC 224

QY 46 SerSerValIysAspArgIleGlyIleThrSerMetIleAlaAspAlaGluGluLeu 65  
DB 225 TCTAGCCGCAAGACAGAGATGTTTATGACATGATTTCTGATGACAGAAAGAGGTCCT 284

QY 66 IleThrProGlyIysSerValIleGluProThrSerGlyAsnThrGlyIleGlu 85  
DB 285 ATCAACACGAGAGAGAGTGTCTATTAAGCCCAAGAGGAACTGGAGTTGGGTTA 344

QY 86 AlaPheMetAlaAlaAlaArgGlyTyrIleLeuIleIleThrMetProAlaSerMetSer 105  
DB 345 GCATTACCGGAGGTGGCAAGGCTTACAGCTTATTAATCAAGCCAGCTTATAGT 404

QY 106 LeuGluArgAlaIleIleLeuLeuAlaPheGlyAlaGluLeuValIleThrAspProAla 125  
DB 405 ACTAGAGAAAGAAATCATTTCTTGTAGCTTTGAGATTGATGGTTTAACTGACCCAGCT 464

QY 126 LysGlyMetIysGlyAlaValAlaGluIleGluIleLeuAlaIleThrProAsnAla 145  
DB 465 AAGGCATGAAAGAGAGCTATGCAAGCCGAGAGATTTGGCGAAACACCCATGCT 524

QY 146 TyrIleLeuGluGluIlePheGluAsnProAlaAsnProIysValHisTyrGluThrGly 165  
DB 525 TACATGCTTCAGCAGTTTGAGAAACCTGCGCAACCTTAAGATCACTATGAGACTACGGGA 584

QY 166 ProGluIleTyrIleThrGlySerAspGlyIleGlyIleAspAlaPheValSerGlyIleGlyThr 185  
DB 585 CCGTAGATATGAAAGACAGCTGCTGGGCAAAATCAATGAGGAGTTTCTTGGATTTGTA 644

QY 186 GlyGlyThrIleThrGlyAlaGlyIleThrLeuIleGluIleGluIleGluIleGlu 205  
DB 645 GGTGATACCATTAAGAGGCTGAGAGATCTTAAGAAACAGAAAGGTCAGCTG 704

QY 206 IleGlyValGluProValGluSerProValLeuSerGlyIleIleProGlyProHisIys 225  
DB 705 TATGAGAGAGAGCCAGTTGAAAGTCTATTCTTATCCGATGGAGAACCGAGTCTCACAG 764

QY 226 IleGluIleGlyIleGlyAlaGlyIlePheIleProGlyValIleGluIleGluIleGlu 245  
DB 765 ATTCAGAGGATGAGAGCTGTTTATACAGATGATTAAGTGTGATCTTATGACGAA 824

QY 246 ValValGluIleSerSerAspGluAlaIleGluThrAlaIleLeuIleGluIleGlu 265  
DB 825 GTTGTTCAGGTTTCAAGATGATGATTCATTCATTCGATGGAGAACCGAGTCTTAAAGAA 884

QY 266 GlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGluIleAlaIys 285  
DB 885 GGTCTTCTTGAGAAATATCATCCGCTGCAAGAGCTGCTGAGCAATTAATCTTGACAG 944

QY 286 ArgProGluAsnAlaGlyIleLeuIleValAlaValPheProSerPheGlyGluArgTyr 305  
DB 945 AGGCCAGAAACCGCTGGGAAAGCTATTGTTGCGATATCCGAGTTTGGGAGAGAGTAT 1004

QY 306 LeuSerSerValIlePheGluSerValArgArgGluAlaGluSerMetThrPheGlu 324  
DB 1005 CTATCAACGATCTTTCATGATCCACAGAAAGAGGAGCCATGACCTTGAG 1061

RESULT 14

ZMCSOATL 1319 bp mRNA linear PLN 28-JAN-2003  
LOCUS ZMCSOATL 1319 bp mRNA linear PLN 28-JAN-2003  
DEFINITION Z.mays mRNA for cysteine synthase.  
ACCESSION X65803  
VERSION X65803.1 GI:758352  
KEYWORDS cysteine synthase; Mcysp gene; O-acetylserine(thiol) lyase.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE  
1 Brander, K.A., Oettlem, G.W. and Brunold, C.  
Isolation of a cDNA (EMBL X65803) encoding a putative chloroplastic isoform of cysteine synthase from maize (PGR95-031)

JOURNAL Plant Physiol. 108, 1748-1748 (1995)  
 REFERENCE 2 (bases 1 to 1315)  
 AUTHORS Brander, K.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAR-1995) K.A. Brander, Institute of Plant  
 Physiology, Altenbergrain 21, CH-3013 Berne, SWITZERLAND  
 FEATURES Location/Qualifiers

source  
 1. 1319  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultiivar="B73 inbred line"  
 /db\_xref="taxon:4577"  
 /tissue\_type="upper 12 root"  
 /dev\_stage="15 weeks"  
 1. 1319  
 /gene="mcy5p"  
 73. 1050  
 /gene="mcy5p"  
 /RC\_number="4.2.99.8"  
 /note="cysteine synthase"  
 /codon\_start=1  
 /product="O-acetylserine (thiol) lyase"  
 /protein\_id="CAA59798.1"  
 /db\_xref="GI:758353"  
 /db\_xref="GOA:P80608"  
 /db\_xref="SWISS-PROT:P80608"  
 /translation="MGEASPSIAKDTVELIGMTPLYTANVTDCYGRRALESMP  
 CSVVKRIIGSMITDAEKLITPPGVALIETPSGNTGIGLAFMAAKGKLTLP  
 SMEMERILIKAFGLAVLTPDPLKMGKAVKABEQAKTPNSYTLAQEPENP  
 YETGPIMVATKIDGLVSGIGTGTITGTGRYLRQPNVYLYGVEVSEVANG  
 GKDPKIKIGTIGAGPILGVLVDLIDETLOVSSDEAIETAKALAKRGLVIGSSGA  
 AAAAVRLAKRPENAGKLFVVFPSPGRYLSVLFQSIKKEASMTVP"  
 73. 180  
 /gene="mcy5p"  
 181. 1047  
 /gene="mcy5p"  
 /product="unnamed"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2 53e-98 Length: 1319  
 Score: 1378.00 Matches: 268  
 Percent Similarity: 91.19% Conservative: 22  
 Best Local Similarity: 84.28% Mismatches: 28  
 Query Match: 84.90% Indels: 0  
 DB: 8 Gaps: 0

US-09-931-457a-31 (1-325) x ZKCSOATL (1-1319)

QY 8 TLeAlAlaSpaValIthrgIuLeuIleGIyVsthrProLeuValIytleuAnlyleu 27  
 Db 94 ATGCCCAAGACGTCACCGAGTGTATGGGAACCGCGCTGTGCTCAACAGAGT 153  
 QY 28 AlAspGIyCySvAlaIaRgValAlaAlaLysleuIuLeuMetGluProCySvSer 47  
 Db 154 ACCGATGGAGGCGCGCCGACGCGCAAGCTCAAGTCCATGAGAGCCCTGCTCAGC 213  
 QY 48 ValIySaPaRgIleGIyYrSerMetIleAlaSpaIaGIuGluLysGIyLeuIleTh 67  
 Db 214 GTCAAGATGATGCTGCTACAGCATGATCAACGACCAAGGAGAAAGGCTGATTA 273  
 QY 68 ProGIyYrSerValIleGIuProThrSerGIyAnthrGIyIleGIyLeuAlaPhe 87  
 Db 274 CCAAGCGTGAAGTGTTCGATTGAACCACTGACGCGCAACAGGCACTGAGCTTT 333  
 QY 88 MetAlaAlaIaRgGIyYrYrLeuIleIleThrMetProIaSerMetSerLeuGIu 107  
 Db 334 ATGCTCTCTGCAAGGCTCAAACTTACACTACATGCTGCTCCTCATGAGAGAG 393  
 QY 108 ArgArgIleIleLeuLeuAlaPheGIyAlaGIuLeuValLeuThraPProAlaLysGIy 127  
 Db 394 AGGAGATCATATTTGAAGGCTTTTGTGTGAACCTGTCTTACTGACCACTCTTGGGA 453

QY 128 MetIyGIyAlaValGIuLysAlaGIuGIuIleLeuAlaLysEThProAnaLalytle 147  
 Db 454 ATGAAAGGAGCTGTCAAGAAAGCGAAGATACAGCAAGACCAACCTCGTACAT 513  
 QY 148 LeuGIuGIuPheGIuAspProAlaAnPProLysValHisTyrGIyThrThrGIyProGIu 167  
 Db 514 CTTCAACATTTGAATATCACTTAACCAAGATTCAGTATGAGCTGATCGTATCGGAGAG 573  
 QY 168 IleTPyGIySerAspGIyYrYrIleAspAlaPheValSerGIyIleGIyThrGIyGIy 187  
 Db 574 ATCTGGAAGCTACAGCAGAGAAATTTGATGCTGTATCTGTATCGTATCGGAGAGAG 633  
 QY 188 ThrIleThrGIyAlaGIyYrYrLeuGIyGIuGIuAnPProAnIleLysLeuIleGIy 207  
 Db 634 ACCATACCGTACTGCGGCAATCTCGAGAGCAAAATCTTAAGTCAAGCTCTATGAT 693  
 QY 208 ValGIuProValGIuSerProValLeuSerGIyGIyYrProGIyProHisLysIleGIu 227  
 Db 694 GTGAGCCAGTGAAGAGTCTGTTTGAATGTGGAAGAACTGACACACAAATTCAA 753  
 QY 228 GIyIleGIyAlaGIyPheIleProGIyValLeuGIuValAsnLeuLeuAspGIuVal 247  
 Db 754 GGAATTTGAGCTGTTTATCCCTGAGCTTGGATGTGATCTCTTATTAATCTCA 813  
 QY 248 GlnIleSerSerAspGIuAlaIleGIuThraLalysLeuAlaLeuLysGIuGIyLeu 267  
 Db 814 CAGCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873  
 QY 268 PheValGIyIleSerSerGIyAlaAlaAlaAlaAlaAlaPheGIuIleAlaLysArgPro 287  
 Db 874 TTGGTTGATCTCTCTCTGTCAGTGAAGTGGCGGCTGTAAGTCTTAAGAGGCCA 933  
 QY 288 GluAsnAlaGIyLysLeuIleValAlaAlaPheProSerPheGIyGluArgTyrLeuSer 307  
 Db 934 GAAAGCCGGAAGATATTGTTGTCGTTCCCGAGTTGCGGAGGCTCAACCTCTCA 993  
 QY 308 SerValLeuPheGIuSerValArgArgIuAlaGIuLysSerMetThrPheGIuPro 325  
 Db 994 TCGTGTCTTCCAGTCACTCAAGAAAGGAGGAGAGATGATGATGATGATGATGATGAT 1047

RESULT 15  
 AF073697 1363 bp mRNA linear PIN 08-APR-1999  
 LOCUS AF073697  
 DEFINITION Oryza sativa cysteine synthase (rcs3) mRNA, complete cds.  
 VERSION AF073697.1 GI:4574138  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Eriatoidae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Nakamura, T., Yamaguchi, Y. and Sano, H.  
 TITLE Four rice genes encoding cysteine synthase: isolation and  
 differential responses to sulfur, nitrogen and light  
 JOURNAL Gene 229 (1-2), 155-161 (1999)

REFERENCE  
 AUTHORS Nakamura, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-1998) Research and Education Center for Genetic  
 Information, Nara Institute of Science and Technology, 8916-5  
 Takeyama, Ikoma, Nara 630-0101, Japan  
 FEATURES Location/Qualifiers

source  
 1. 1363  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultiivar="Yipponbare"  
 /db\_xref="taxon:33947"  
 1. 1363  
 /gene="rcs3"

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rge

Page 15

CDS

```
139. .1116
/gene="rcs3"
/EC number="4.2.99.8"
/note="O-acetylserine (thiol)-lyase"
/codon_start=1
/product="cysteine synthase"
/protein_id="ABD23903.1"
/db_xref="GI:4574139"
/translation="MAESQSIASDVTALIGNTPLVINKVVDGCEAIAKLBIMEP
SSVVDRIEAGSWITDAERKGLITPGKSLIETSGNTGIGLAFMAAGYCLITMPA
SMSEMERILKAFGLVLTDPILGMKALIOKADLAKMENSYTLQFENPAPKIH
YETGPETWKATAGKVDILVSGIGTGTGKIKONKEIKIYGEYPESAILSG
GRPGPHKIQIGAGFVPLVDVNLIDRVVQUSMDPAISMAKQIALKSLIYGISGAA
AAVAIRVAQRPEKGLVYVVPFSGERTLSVLPESIKRANMVP"

```

ORIGIN

Alignment Scores: 3.14e-98 length: 1363  
Pred. No.: 1377.00 Matches: 263  
Score: 89.54% Conservative: 28  
Percent Similarity: 80.92% Mismatches: 34  
Best Local Similarity: 84.84% Indels: 0  
Query Match: 8 Gaps: 0  
DB: 8

US-09-931-457a-31 (1-325) x AF073697 (1-1363)

```
QY 1 MetAlaValGluArgSerGlyIleAlaIysAspValThrGluLeuIleGlyIleThrPro 20
DB 139 ATGGCGAATCGAGCCCAATCGATTCGAGTGAATGACTGCATGATCGCAATCTCC 198
QY 21 LeuValIleLeuAsnIleLeuAlaAspGlyCysValAlaArgValAlaAlaIysLeuGlu 40
DB 199 CTGGTCTACTTAAACAAGTGTGATGTGATGTGAGCCCAATCGCTGCCAAGCTTGAG 258
QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyIleThrSerMetIleAlaAspAla 60
DB 259 ATCATGAGAGCCCTGCTCCAGTGTCAAGATGATGATGATGATGATGATGATGATGATG 318
QY 61 GluGluArgGlyLeuIleThrProGlyIleSerValIleLeuIleGlyIleThrSerGlyAsn 80
DB 319 GAAAGAAAGGCTCTCACTCACTCAAGAAAGAGGTTTGAATTAAGCTCTCACTCACTCACT 378
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyIleThrIleLeuIleThrMet 100
DB 379 ACTGGATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 438
QY 101 ProAlaSerMetSerLeuGluIleArgGlyIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
DB 439 CCAGATTCATAGCACTGGAAGAGAGATCTTAAAGCTTCTGAGTCCAGATTTGATC 498
QY 121 LeuThrAspProAlaIysGlyMetIysGlyAlaValAlaGluIleAspGluIleLeuAla 140
DB 499 CTTAACGATCCACTTCTGAGCATGAAAGGTGCATCCAGAGCAGATGAGCTGGCAGCA 558
QY 141 LysThrProAsnAlaTyrIleLeuGluIlePheGluIleAsnProAlaAsnProIysValHis 160
DB 559 AAGATGCCATTAACCTTCACTTCTCCAGCAGTTGAGAACCTTGCACCAACCCAAAGATCCAC 618
QY 161 TyrGluThrThrGlyIleProGluIleTyrPlySerAspGlyIleValIleAspAlaPheVal 180
DB 619 TATGAGACACTGAGCTGAGATTTGAGAGGCCACTGCGGAAAGTTGATCTCTTGTTC 678
QY 181 SerGlyIleGlyIleThrGlyIleThrIleThrGlyAlaGlyIleValIleLeuIleGluIleAsn 200
DB 679 TCTGGATTTGGAACCGTGTGTACAGTAAACCGAAGCAGAAAGTACTCTCAAGAGCAGAAC 738
QY 201 ProAsnIleLeuLeuIleGlyIleValGluIleProValGluSerProValIleSerGlyIleVal 220
DB 739 CCTGAGATTAAGATCTATGATGTTGAAACCGAAGAGTCAATTATCTGAGAGAGA 798
QY 221 ProGlyIleProIleIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGluIleVal 240
```

```
QY 241 AsnLeuLeuAspGluValValGluIleSerSerPheGluAlaIleGluThrAlaIysLeu 260
DB 859 AATCTCTTGAATGAAAGGCTCCAGTGTCAAGTATTAAGCCATCAATGCGAAGCAG 918
QY 261 LeuAlaLeuIysGluGlyIlePheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280
DB 919 CTACGCTTGAAGAAAGGCTCTGTTGGATCTCATCTGGTGTGCTGAGAGTGGCTGCC 978
QY 281 PheGluIleAlaIysArgProGluAsnAlaGlyIleLeuIleValAlaAlaPheProSer 300
DB 979 ATCAGGCTCGCGCAGCGCCCGAGAACAAAGGAAAGCTGTGAGTTGTGCTTCCGAGC 1038
QY 301 PheGlyIleGluArgTyrIleSerSerValIlePheGluSerValArgArgGluAlaGluSer 320
DB 1039 TTGGTGAAGCCTTACTCTGTCTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1098
QY 321 MetThrPheGluPro 325
DB 1099 ATGGTGTTCAGAGCA 1113
```

Search completed: May 21, 2004, 23:43:15  
Job time : 4706 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: May 21, 2004, 21:58:57 ; Search time 3100 Seconds

(without alignments)  
3130.712 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIADVTTELIGTP.....LSVLFESVREASMTFEP 325

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgrt 1/USPO.spool/us09931457/runat.18052004.121729.20307/app.query.fasta\_1.519  
-DB=EST -QFMT=fastap -SUPPLX=p2n.rst MINMATCH=0.1 -LOOFCU=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=ptc -NORM=ext -HRAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us09931457 @CGN 1 1 1906 @runat.18052004.121729.20307 -NCPU=6 -ICPU=3  
-NO MAP -IARGOUBERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WAW TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FAPOP=6  
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estdpl:\*  
7: em\_estdro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estlan:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrc:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_dhg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1417	87.3	994	14	CK260011	CK260011 EST706089
2	1392	85.8	929	14	CF513886	CF513886 Cabud007
3	1389	85.6	1483	11	AY103651	AY103651 Zea mays
4	1372	84.5	967	14	CK271210	CK271210 EST717288
5	1367	84.2	910	14	CB974516	CB974516 CAB30005
6	1364	84.0	937	14	CK266154	CK266154 EST712232
7	1313	80.9	881	14	CK259489	CK259489 EST705567
8	1311	80.8	907	14	CK260012	CK260012 EST706090
9	1311	80.8	920	14	CK259857	CK259857 EST705935
10	1310	80.7	873	14	CK259488	CK259488 EST705936
11	1306	80.5	917	14	CK259858	CK259858 EST725584
12	1291	79.5	935	14	CK279506	CK279506 EST705567
13	1285	79.2	934	29	CG848066	CG848066 ZMGBB031
14	1269	78.2	1288	11	AY112530	AY112530 Zea mays
15	1260	77.6	886	14	CK265667	CK265667 EST711745
16	1250	77.0	772	14	CA801406	CA801406 esau05c03
17	1221	75.2	802	14	CB974696	CB974696 CAB30005
18	1214	74.8	845	14	CF200608	CF200608 R8R90915N
19	1214	74.8	1155	14	CF162330	CF162330 FGAS01492
20	1206	74.3	792	14	CF512887	CF512887 Cabud0005
21	1197	73.8	965	14	CK274117	CK274117 EST720195
22	1194.5	73.6	803	12	BG648907	BG648907 EST510526
23	1192	73.4	1111	14	CK163592	CK163592 FGAS01622
24	1186	73.1	724	10	BF003458	BF003458 EST431956
25	1186	73.1	813	14	CF205785	CF205785 R8R909151
26	1174	72.3	791	12	BG595822	BG595822 EST494500
27	1172	72.2	888	14	CF526675	CF526675 EST712824
28	1156	71.2	856	14	CF526675	CF526675 EST712824
29	1155	71.2	914	14	CD439033	CD439033 R010N0520
30	1149	70.8	730	13	BU027058	BU027058 QG18P11
31	1138	70.1	785	12	BG587892	BG587892 EST489667
32	1117	68.8	866	14	CD435997	CD435997 B010N0369
33	1110	68.4	673	12	BM815595	BM815595 EST593689
34	1107	68.2	734	12	BG648870	BG648870 EST510489
35	1104	68.0	726	12	BG128810	BG128810 EST74456
36	1088	67.0	742	14	CD813753	CD813753 BNT5. 0200
37	1084	66.8	655	12	B1309270	B1309270 EST530680
38	1081	66.6	740	13	BQ914877	BQ914877 QHB12F18
39	1075	66.2	804	14	CB658211	CB658211 OSJNEC14B
40	1073	66.1	694	14	CK273767	CK273767 EST719845
41	1071	66.0	686	12	B1264323	B1264323 NF118G04P
42	1070	65.9	678	10	AW311624	AW311624 SGT43B07 Y
43	1068	65.8	800	14	CB651906	CB651906 OSJNEC01B
44	1066	65.7	806	14	CB654896	CB654896 OSJNEC07M
45	1064	65.6	768	14	CD824490	CD824490 BNT5. 0531L

ALIGNMENTS

RESULT 1  
LOCUS CK260011  
DEFINITION EST706089 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB228 5' end, mRNA sequence.  
ACCESSION CK260011  
VERSION CK260011.1 GI:39816989  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.



## REFERENCE

1 (bases 1 to 994)  
 AUTHORS Buell, C.R., Hart, A., Zisman, V., Karanicheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from abiotic stressed potato tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST706088 EST706090 EST706091  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr. Rockville, MD 20850, USA  
 Email: potato-array@igir.org  
 Clones can be requested from TIGR via potatoest@tigr.org  
 Seq primer: ATG TAG GTC ACA CTA TAG.

## FEATURES

source

1. .994  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /culivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="PDB228"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Prod. No.: 5,44e-143 Length: 994  
 Score: 1417.00 Matches: 279  
 Percent Similarity: 94.10% Conservative: 24  
 Best Local Similarity: 86.65% Mismatches: 19  
 Query Match: 87.31% Indels: 1  
 DB: 14 Gaps: 0

US-09-931-457A-31 (1-325) x CK260011 (1-994)

QY 1 MetAlaValGluATGAGSerGlyTleAlaAlaValAspValThrGluLeuIleGlyLysThrPro 20  
 Db 28 ATGCGCGGCGGAAAGATTGGAATGCCAGATGTAATGATGATGTAACACTCT 87  
 QY 21 LeuValTyrIleuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40  
 Db 88 TTGGTATACCTGGAATATGTTGATGCGATGGGTGTGACCTGTGTGCTGCCAACTGGAA 147  
 QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 148 AGCATGAGCCATGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 207  
 QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
 Db 208 GAGGAGAGAGGCTTCATCAACACCGGCGAGAGTGTCTCCATCGAAGCAAGGAGAAAC 267  
 QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100

Db 268 ACTGGTGTGAGTGGCATTCATGCTGCTGCTGAAGGCTTAACTCATTCATTCAGATG 327  
 QY 101 ProAlaSerMetSerLeuGluIubArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 Db 328 CCTCTTCAATGATGCTTGAAGAGAAATTAATTCGCGCTTCGCTGCTGATGGTG 387  
 QY 121 LeuThrAspProAlaLysGlyMetLysGlyValAlaGlnLysAlaGluIleLeuAla 140  
 Db 388 CTACCGATTCAGCAAAAGGATTAAGGCTCTATTCAAAAGCTTAAGATTAAGGCC 447  
 QY 141 LysThrProAsnAlaTyrIleLeuGluGlnPheGluAsnProAlaAsnProLysValHis 160  
 Db 448 AAAACACCACTCATTAATTCCTTCAGCAATTTGAAAACCTGCTAACCCAAAGATACAC 507  
 QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
 Db 508 TATGAGACCACTGCTCTAGATCTGGAAGGCTCAAAAGGCTCAAAAGGATGCTCTAGTC 567  
 QY 181 SerGlyIleGlyThrGlyLysThrIleThrGlyAlaGlyLysTyrLeuLysGluAsn 200  
 Db 568 TCTGCAATTCGAACAGAGGACAGATTAATCTGCTCAGGCAAGATTTGAGAGCAGAAC 627  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
 Db 628 CCCAAGCTTAAGCTTATGAGCGGTGAACAGATTAAGTGAAGCTATCTTCTGCTGAAG 687  
 QY 221 ProGlyProHisIleLysIleGluGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 Db 688 CTCGGTCCACATTAAGTTCAAGGAGATTTGCTGCTGCTTCTTCTGCTGCTTGAAGTT 747  
 QY 241 AsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 Db 748 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyValAlaAlaAlaAla 280  
 Db 808 CTGCAATTAAGAGAGATTCATGAGGAGATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 867  
 QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPheProSer 300  
 Db 868 ATTAAGATGCTTAAGCGCTGGAAGATTCGGAAGCTCATTTGTTAT-TTTCACAC 926  
 QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320  
 Db 927 TTGGAGAGCGATATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986  
 QY 321 MetThr 322  
 Db 987 ATGACT 992

RESULT 2  
 LOCUS CF513886  
 DEFINITION Bud - CABUD Vitis vinifera cDNA clone Cabud0007\_11if\_H10 5', mRNA  
 ACCESSION CF513886  
 VERSION CF513886  
 KEYWORDS EST  
 SOURCE Vitis vinifera  
 ORGANISM Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosids; Vitaceae; Vitis.  
 1 (bases 1 to 929)  
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and Cook, D.  
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages  
 Unpublished (2003)  
 CONTACT: Douglas Cook, PhD  
 CABS Genome Facility

UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcoc@ucdavis.edu  
Seq primer: ACGGTACCGACATATGCC.

## FEATURES

source

location/Qualifiers

1..929

/organism="Vitis vinifera"

/mol\_type="rRNA"

/cultivar="Cabernet Sauvignon (Clone 8)"

/db\_xref="taxon:29760"

/clone="CABUD007.11IF\_H10"

/sex="Hermaphrodite"

/dev\_stage="pre-bloom (10-11 days before bloom)"

/lab\_host="DM5alpha"

/clone\_id="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"

/note="Organ: Bud; Vector: pDNR; Site: 1: Sfil; Site: 2: Sfil; CABUD is a cDNA library of Vitis vinifera cv. Cabernet Sauvignon, Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:

5'-AAGCAGGTATACACGAGGAGGCGCATTAAGCGCGG-3' and 5'-ATCTAGAGCGGAGGCGCGCGGAGCATG-3' (30NN-3'). Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,53e-140 Length: 929  
Score: 1392.00 Matches: 269  
Percent Similarity: 95.74% Conservative: 23  
Best Local Similarity: 88.20% Mismatches: 13  
Query Match: 85.77% Indels: 0  
DB: 14 Gaps: 0

US-09-931-457A-31 (1-325) x CFS13886 (1-929)

OY 21 LeuVallyrLeuAsnHlyLeuAlaAspGlyCyValAlaArgValAlaAlaAlaLeuGlu 40  
Db 1 TTAGTATATCTGAAACATATGCTGATATGCTGATGCTGATGCTGCAAACTTGTAA 60  
OY 41 LeuMetGluProCySerSerVallyAspArgIleGlyTYrSerMetIleAlaAspAla 60  
Db 61 ATGATGAGGCTTGTCTAGTGTCTAGTGTCTAGTGTCTAGTGTCTAGTGTCTAGTGTCA 120  
OY 61 GluGluysGlyLeuIleThrProGlyIlySerValIleuIleGluProThrSerGlyAsn 80  
Db 121 GAAGAGAGGAGGCTCATTAAGCGGAGGAGGAGTGTCCATATGAGCTTACCAAGTGTAT 180  
OY 81 ThrGlyIleGlyLeuAlaPheMetIleAlaAlaArgIlyTYrIlySerMetIleThrMet 100  
Db 181 ACTGGCATAGGAGGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
OY 101 ProAlaSerMetSerIleuGlyArgIleIleLeuAlaPheGlyAlaGluLeuVal 120  
Db 241 CCGTGTTCGATGAGTGTGAGT 300  
OY 121 LeuThrAppProAlaIlyGlyMetIlySerGlyAlaValGlyIlyAsnGluGluIleuVal 140  
Db 301 CTCACAGATCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
OY 141 LysThrProAsnAlaTYrIleuGluGluIleuPheGluAsnProAlaAsnProIlyValHis 160  
Db 361 AAGACACCTTATCTCTATATCTTCAAGCAATTTGAAACCTGCGCAATCCAGATTCAT 420  
OY 161 TyrGluThrThrGlyProGluIleThrIlySerIlySerArgGlyIlyValAspAlaPheVal 180

Db 421 TATGAACCACTGGACCGAGACTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
OY 181 SerGlyIleGlyThrGlyValTYrIleThrGlyAlaIleIlySerTYrLeuIlyGluIleAsn 200  
Db 481 TCTGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
OY 201 ProAsnIleIlySerIleGlyValAlaGluProValGluIlySerProValIleuSerGlyIly 220  
Db 541 TCTGACATTAAGCTGTATGATGTATGATGATGATGATGATGATGATGATGATGATGATG 600  
OY 221 ProGlyProIlyIlyIleGlyIlyIleGlyAlaGlyIlyPheIleProGlyValLeuGluVal 240  
Db 601 CCGTGGCGGATTAAGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
OY 241 AsnLeuIlyAspGluValAlaGluIleSerSerAspGluIlyIleGluIlyIleValAlaIle 260  
Db 661 AATTTGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
OY 261 LeuAlaLeuIlyGlyIlyLeuPheValIlyIlySerSerGlyAlaAlaAlaAlaAlaAla 280  
Db 721 CTTCGATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
OY 281 PheGluIleAlaIlyAsnArgProGluIlyAsnAlaGlyIlyIlyLeuIleValAlaPheProSer 300  
Db 781 ATTAAGTTAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
OY 301 PheGlyIlyIlyIlyIlyLeuSerSerValIlyPheGluIlySerValArgArgIlyIlySer 320  
Db 841 TTTCGAGAGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
OY 321 MetThrPheGluPro 325  
Db 901 ATGCTGTTGAACCC 915

## RESULT 3

AY103651

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITL

JOURNAL

REFERENCE

AUTHORS

TITL

JOURNAL

REFERENCE

AUTHORS

TITL

JOURNAL

REFERENCE

AUTHORS

TITL

JOURNAL

REFERENCE

AUTHORS

TITL

JOURNAL

REFERENCE

AUTHORS

TITL

AY103651 Zea mays P00145346 mRNA sequence.  
AY103651 Zea mays P00145346 mRNA sequence.  
AY103651.1 GI:21206729  
HTC.  
Zea mays  
Zea mays  
Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 1483)  
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1483)  
Coe, E.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZMDB and may be found by BLAST  
searching at WBL, maizeidb.org; ZMDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZMDB:  
www.zmdb.iastate.edu.  
location/Qualifiers  
1..1483  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:638956"  
/db\_xref="taxon:4577"  
/clone\_id="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST







QY 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160  
 DB 440 AANAACCCCAACCCCTATATCTTCAGCATTTGAAACCCCTGTAACCCAAAGATACAC 499  
 QY 161 TyrGlnThrThrGlyProGluIleTyrGlySerAspGlyLysIleAspAlaPheVal 180  
 DB 500 TATGAGACCACTGCTCTGAGATCTGAAAGGCTCAAAAGGATGAGATGCTCTAGTC 559  
 QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlnGlnAsn 200  
 DB 560 TCTGGAATTTGAAACGAGGACGATACCTGCTTCAGGACGATTTTGAAGGACGAGAC 619  
 QY 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
 DB 620 CCAACACTTAAGCTGTATGCGCGTGAACCAAGTGAAGTCTATCTTCTGCTGGAAG 679  
 QY 221 ProGlyProHisLysIleGlnGlyIleGlyValGlyPheIleProGlyValLeuGluVal 240  
 DB 680 CTTGGTCCATTAAGATTCAGGAGATTTGCTGCTGCTTCTCTGCTGCTGCTGCTGCT 739  
 QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 DB 740 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799  
 QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyLysAlaAlaAlaAla 280  
 DB 800 CTGGCATTTGAAAGAGATGCTAGTGGGATTTCTATGCTGCTGCTGCTGCTGCTGCT 859  
 QY 281 PheGlnIleAlaLysAspProGluAsnAlaGlyLysLeuIleValAlaLysPheProSer 300  
 DB 860 ATTAAGTCTTAAGGCGCCCTGAAGTCTGAAGTCTATTTTGTGATTTTTCACAG 919  
 QY 301 PheGlyGluArgTyrLeu 306  
 DB 920 TTCGAGAGCGATATCTT 937  
 RESULT 7  
 CK259489 881 bp mRNA linear EST 12-DEC-2003  
 LOCUS EST705567 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 DEFINITION  
 CK259489  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 881)  
 Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A. and Baker, B.  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 JOURNAL  
 TITLE  
 AUTHORS  
 COMMENT  
 Other ESTs: EST705565 EST705566  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: CAG GAA ACA GCT ATG ACC.  
 FEATURES  
 source  
 1. 881  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAB060"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Toma"  
 /clone\_lib="Potato abiotic stress cDNA library"  
 /note="Vector: PCWSPort6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25

C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at the following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, 2d, and 4d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d, and 4d were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Prod. No.: 8,89e-132 Length: 881  
 Score: 1313.00 Matches: 253  
 Percent Similarity: 94.528 Conservative: 22  
 Best Local Similarity: 86.648 Mismatches: 16  
 Query Match: 80.908 Indels: 0  
 DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CK259489 (1-881)

QY 4 GluArgSerGlyTyrIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyr 23  
 DB 2 GAAGAATTGGATTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 61  
 QY 24 LeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuLysLeuMetGlu 43  
 DB 62 CTGAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121  
 QY 44 ProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLys 63  
 DB 122 CCATGCTCTAGGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181  
 QY 64 GlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIle 83  
 DB 182 GATTCATTAACCCCGGAGAGATGCTCTCAATCAAGCTCAAGATGAGAAACCTGCTGA 241  
 QY 84 GlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProLysSer 103  
 DB 242 GGAATGGCATTAAGCTGCTCTCAAGGCTCAAACTCATATTAAGATGATGATGATGATGAT 301  
 QY 104 MetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAsp 123  
 DB 302 ATGAGCTTGAAGAGAAATTAATTCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 361  
 QY 124 ProAlaLysGlyMetLysGlyValAlaGlnLysAlaGlnLysIleLeuAlaLysThrPro 143  
 DB 362 CCAAGAAAGGATGAGAGGCTCTATTTCAAGGCTGAAGATGAAGATGAAGGCAAAACACC 421  
 QY 144 AsnAlaTyrIleLeuGlnGlnPheGluLeuAsnProAlaAsnProLysValHisTyrGluThr 163  
 DB 422 AACCTCTAATTTCTGAGATTTAAACCCCTGCTCAACCCAAAGATCACTATGAGACC 481  
 QY 164 ThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIle 183  
 DB 482 ATGGTCTCGATGATGAGAGGCTCAATATGGAAGATGAGATGCTCTGATGAT 541  
 QY 184 GlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGluAspProAsnIle 203  
 DB 542 GGAACGAGAGGACATTAAGTTCAGGCAAGATTTTGAAGAGACCAAGCCCAACGCT 601  
 QY 204 LysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGlyPro 223





Db 871 ATTAAGTCCTAAGCCCTGAAATGCTGGAG 906

RESULT 9  
CK259857 920 bp mRNA linear EST 12-DEC-2003  
LOCUS EST05935 potato abiotic stress cDNA library Solanum tuberosum cDNA  
DEFINITION clone POAB178 5' end, mRNA sequence.  
CK259857  
ACCESSION CK259857.1 GI:39816835  
VERSION EST.  
KEYWORDS Solanum tuberosum (potato)  
SOURCE Solanum tuberosum  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 920)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST705934 EST705936 EST705937  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..920  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB178"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCKVSPORT6.1, Site 1: EcoRI, Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr;  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

FEATURES  
source

ORIGIN

Alignment Scores:

Pred. No.: 1.58e-131 Length: 920  
Score: 1311.00 Matches: 253  
Percent Similarity: 94.18% Conservative: 22  
Best Local Similarity: 86.64% Mismatches: 17  
Query Match: 80.78% Indels: 0  
DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CK259857 (1-920)

QY 1 MetAlaValGluArgSerGlyLeuAlaValAspValThrGluLeuGlyLysThrPro 20  
|||||  
|||||

Db 45 ATGGCGGGGAAAAAGATTGGAATGCCAGATGTACTGAATGATTGATGACATCTCT 104  
QY 21 LeuValTyrLeuLeuLeuValAspGlyCysValAlaArgValAlaAlaLeuLeuGlu 40  
|||||  
105 TTGGTATACCTGGAATTAATGTTGGATGGGATGGTGGTACAGTGTGCTGCCAGCTGGA 164  
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
165 AGCATGAGGCGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224  
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
225 GAGGAGAGAGGATCTCAACCAACCGGAGAGAGTCTCTCATCGAATCAGAGAGAGAGAAC 284  
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaArgValTyrLysLeuIleIleThrMet 100  
285 ACTGGTATGAGATTGGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 344  
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
345 CTTCTTCATGATGATCTTGAGAGAGAGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 404  
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValGlnLysAlaGluIleLeuAla 140  
405 CTTACCATTCAGCAAG 464  
QY 141 LysThrProAlaAlaTyrIleLeuGlnIlePheGluLeuProAlaAspProLysValHis 160  
465 AAACACCAACCACTCTATCTCTCAACCAATTTGAAAAACCTGCTAACCCAAAGATCAC 524  
QY 161 TyrGlnThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180  
525 TATAGACCACTGATCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584  
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200  
585 TCTGGAATTTGGAACAGAGAGAGAGAGAGAGATTAATCTGCTGCTGCTGCTGCTGCTGCT 644  
QY 201 ProAlaIleLysLeuIleGlyValGluProValGluSerProValLysSerGlyLys 220  
645 CCCACCACTTAAGCTGATGAG 704  
QY 221 ProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
705 CTTGCTCAACATTAAGATTCAGGGGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 764  
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
765 AACCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824  
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAlaAla 280  
825 CTGCGATTGAGAGAGAGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884  
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLys 292  
885 ATTAAGTCCTAAGCCCTGAAATGCTGGAG 920  
Db

RESULT 10  
LOCUS CK259488 873 bp mRNA linear EST 12-DEC-2003  
DEFINITION EST705566 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB060 5' end, mRNA sequence.  
ACCESSION CK259488  
VERSION CK259488.1 GI:39816466  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 873)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.



TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST705565 EST705567  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG CTG ACA CTA TAG.  
Location/Qualifiers

## FEATURES

source

1. 873  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POA8060"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-T01A"  
/clone\_id="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,  
2d. Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.86e-131 Length: 873  
Score: 1310.00 Matches: 252  
Percent Similarity: 94.50% Conservative: 23  
Best Local Similarity: 86.60% Mismatches: 16  
Query Match: 80.71% Indels: 0  
DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CK259488 (1-873)

QY 3 ValGluArgSerGlyIleAlaIleAspValThrGluLeuIleGlyLysThrProLeuVal 22  
DB 1 ATGAGAAAAGATGGAATGTCAGAGAGTAACTGAATGATGGTAAACACCTTTGGTA 60  
QY 23 TyrluLeuIleAlaIleAspGlyCysValAlaIleArgValAlaAlaIleLeuGluLeuMet 42  
DB 61 TACCGAATTAATGTTGAGATGGGAGTGTGACATGTTGCTGCTCCACCTGGAAGCATG 120  
QY 43 GluProCysSerSerValIleAspArgIleGlyLysSerMetIleAlaAspAlaGluGlu 62  
DB 121 GAGCATGCTCTAGAGTTAAGATATGATGATATAGTATGATTACAGATGCTGAGAG 180  
QY 63 LysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAspThrGly 82  
DB 181 AAGGCTTCATCAAAACCCGGGCGAGGTGTCCTCATCAAGAACTTACAGTGGAAACATGGT 240  
QY 83 IleGlyLeuAlaPheMetAlaAlaIleArgGlyTyrLysLeuIleIleThrMetProAla 102  
DB 241 GTAGCATTTGCAATTCATGCTGCTCTAAAGGCTCAACATCATCATTCATGATGCTTCT 300

QY 103 SerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValAlaGluLeuValLeuThr 122  
DB 301 TCAATGAGCTTGAAGAGAAATATTTCTGGCTGCTTCGGTGCTGAGTGGTCTTACC 360  
QY 123 AppProAlaIleGlyMetIleGlyValAlaGluMetIleGluIleLeuAlaIleThr 142  
DB 361 GATCCAGCAAAAGGATGAGGATGAGGATCTTTCAAGAGGCTGAAGATTAAGGCCAAACAA 420  
QY 143 ProAsnAlaTyrIleLeuGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 162  
DB 421 CCCAATCTCTAATTTCTTACGCAATTTGAAACCCCTGTCAACCAAGATACATCATGAG 480  
QY 163 ThrThrGlyProGluIleThrLysGlySerAspGlyLysIleAspAlaPheValSerGly 182  
DB 481 ACCACTGCTCTGAGATCTGGAAGGCTCAATGGAAGGATGATGCTTACTTCTTGA 540  
QY 183 IleGlyThrGlyGlyThrThrLeuAlaGlyLysTyrLysLeuLysGluIleAsnProAsn 202  
DB 541 ATTGGAACAGAGGACCAATATCTGTTACGCAATTTTGAAGAGCAAGACCCCAAC 600  
QY 203 IleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLysProGly 222  
DB 601 GTTAGCTGTATGCGGTGGAACCACTGGAATGCTATCTTCTGTTGGAAGCCTGT 660  
QY 223 ProHleValIleGlnGlyLysIleGlyLysIleGlyLysIleGlyLysIleGlyLysIle 242  
DB 661 CCACATTAAGATTCAGGAGATGAGTGTGCTGCTTCAATCTGAGTTTGAAGATTAACCT 720  
QY 243 LeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeuAla 262  
DB 721 ATTAGAGATGTAGTTAGTTTCAAGTGAAGATCAATGAATGCTAAGCTTCTGGA 780  
QY 263 LeuLysGluGlyLeuPheValGlyLysSerSerGlyAlaAlaAlaAlaPheGln 282  
DB 781 TTGAGAGGAAGATGTGATGTGAATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 283 IleAlaLysArgProGluLysAlaGlyLysLeu 293  
DB 841 GTGCTTAAGCCCTGAAATGCTGGAAGCTC 873

## RESULT 11

CK259858

LOCUS

DEFINITION

CK259858 917 bp mRNA linear EST 12-DEC-2003  
EST705936 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POA8178 5' end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

SOLANUM

TUBEROSUM

## FEATURES

source

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished (2003)  
Other ESTs: EST705934 EST705935 EST705937  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: CAG GAA ACA GCT ATG ACC.  
Location/Qualifiers  
1. 917  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POA8178"  
/tissue\_type="abiotic stress treated leaf and root tissue"

/lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /note="Vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,  
 2d, and 4d). Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cDNA library. RNA sample."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,47e-131 Length: 917  
 Score: 1306.00 Matches: 252  
 Percent Similarity: 94.16% Conservative: 22  
 Best Local Similarity: 86.60% Mismatches: 17  
 Query Match: 80.47% Indels: 0  
 DB: 14 Gaps: 0

US-09-931-457a-31 (1-325) x CK259858 (1-917)

Qy 1 MetAlaValGuaTgSerGlyTleAlaAlaValAspValThrgluLeuIleGlyLeuThrPro 20  
 Db 45 ATGCGCGCGGGAAGAAGATTGGATTCGCAAGAGTAACTGATTTGATTGGTAACATCTCT 104  
 Qy 21 LeuValIyrieuAnuLysLeuAlaAspGlyCyValAlaAlaArgValAlaAlaValLeuGlu 40  
 Db 105 TTGGTAAACCTGGAATTAATGTTGGATGCGGTGCTGTCACGCTGTGCTGCCAACCTGAA 164  
 Qy 41 LeuMetGluProCySerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60  
 Db 165 AGCATGAGCCATCTCTAGCTTAAAGATACGATCGCTTAACTAATGATTAACGATGCT 224  
 Qy 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80  
 Db 225 GAGGAGAAGGGGTCTCATCAACCCGGCGAGAGTCTCTCATCAAGCTAAGAGTGGAAC 284  
 Qy 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100  
 Db 285 ACTGGTGTAGGATGGCATTCATGCTGCTCTAAAGGCTCAACAACCTCAATCATTCAGATG 344  
 Qy 101 ProAlaSerMetSerLeuGluArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 Db 345 CCTCTTCAATGAGCTTGAAGAGAATATTCTCGCGTCTTGGTGTGCTGAGATTTGCTG 404  
 Qy 121 LeuThrAspProAlaIleGlyMetLysGlyAlaValGluLysAlaGluGluIleLeuAla 140  
 Db 405 CTTAACCATTCACGAAAGGATGAGAGGCTCTTAATTCAAAGGCTGAAGAGATTAAGGCC 464  
 Qy 141 LysThrProAlaAlaIleLeuGluGlnPheGluLysProAlaAsnProLysValHis 160  
 Db 465 AAAACACCAACTCTATTTCTTCAAGAAATTTGAAAACTGTGTAACCCAAAGATACAC 524  
 Qy 161 TyrGluThrThrglyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180  
 Db 525 TATGAGCACTGCTGCTGAGATCTGAAAGGCTCAATATGGGAAAGTGAGATCTTATGTC 584

Qy 181 SerGlyIleGlyThrglyThrIleThrglyAlaGlyLysTyrLeuLysGluGluAsn 200  
 Db 585 TCTGGAATTTGGAACAGAGAGCAGAAATACCTGCTGGCGCAAGTATTGAGAGAGCAAG 644  
 Qy 201 ProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGlyLys 220  
 Db 645 CCGACGTTAGCTGATATGCGGTGAGAACAGTTGAAGTCTATCTCTTCTGGGGAAG 704  
 Qy 221 ProGlyProHisLysIleGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240  
 Db 705 CCGGTTCACATTAAGATTGAGGGAATGGGTGCTGCTTCTTCCTGCTGTTTGAAGTT 764  
 Qy 241 AsnLeuLeuAspGluValValGluIleSerSerAspGluAlaIleGluThrAlaLysLeu 260  
 Db 765 AACCTTATATATATGATTTGCTAGGCTTCAAGTGAAGATTCATTAAGATGCTTACCT 824  
 Qy 261 LeuAlaLeuLysGluGluLysPheValGlyIleSerSerGlyValAlaAlaAlaAla 280  
 Db 825 CTGGCATTTGAAGAGAGATTGCTAGTGGGAAATTTCTGCTCTGCGCGCGGCA 884  
 Qy 281 PheGluIleAlaLysArgProGluAsnAlaGly 291  
 Db 885 ATTAAGTCTGTAAGCGCCCTGAATAAGCTGGG 917

RESULT 12  
 LOCUS CK279506 935 bp mRNA linear EST 12-DEC-2003  
 DEFINITION BSE725584 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 ACCESSION CK279506  
 VERSION CK279506.1 GI:39836484  
 KEYWORDS EST,  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum

REFERENCE  
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from abiotic stressed potato tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_ESTs: BSE725585  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ACT TAG GCG ACA CTA TAG.

## FEATURES

## Source

1..935  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAEC59"  
 /tissue="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /note="Vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,  
 2d, and 4d). Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were heat stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Prod. No.:	2 41e-129	Length:	935
Score:	1281.00	Matches:	254
Percent Similarity:	93.27%	Conservative:	23
Best Local Similarity:	85.52%	Mismatches:	20
Query Match:	79.54%	Indels:	1
DB:	14	Gaps:	0

US-09-931-457a-31 (1-325) x CK279506 (1-935)

```

QY 1 MetAlaValAluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
DB 46 ATGCGCGGGGAGAAAGATTGGATCCGACGATGTAATGATTGTTGACCTCTCT 105
QY 21 LeuValIleuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 106 TTGGTATACCTGTAATATGTTGGAGATGGATGGATGCTGCTGCTGCTGCAAG 165
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyLysSerMetIleAlaAspAla 60
DB 166 AGCATGAGGACCATGCTAGTGTATGAAGATGAGATCGGTATATGATGATTACAGATGCT 225
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
DB 226 GAGAGAGAGAGGCTCATCAAAACCGGCGAGAGATGCTCTCACTCAAGCTTACAGTGAAC 285
QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgIleGlyLysLeuIleIleThrMet 100
DB 286 ACTGCTGATGAGATTCGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
QY 101 ProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
DB 346 CCTCTTCATGAGCTCTGAGAGAGAAATATTCGCTGCTGCTGCTGCTGCTGCTGCTG 405
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGluLeuIleLeuAla 140
DB 406 CTTACCGATCCAGCAAAAGGAGATGAGGCTGCTATTCAAAGGCTGAGAGATGAAGGCC 465
QY 141 LysThrProAsnAlaIleIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
DB 466 AAAACACCAACCTCTATATCTTCAGCAATTTGAAAACCTGCTAACCCAAAGATACAC 525
QY 161 TyrGluThrThrGlyProGluIleIleThrLysGlySerAspGlyLysIleAspAlaPheVal 180
DB 526 TATGAGACCATGCTGCTGAGATTCGAGAAAGGCTCAAAATGGAAGGATGCTGATGCT 585
QY 181 SerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysGlyLeuLysGluGlnAsn 200
DB 586 TCTGGAATTCGAAACGAGGACACGATTAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 201 ProAsnIleLysLeuIleGlyValAlaGluProValGluSerProValLeuSerGlyLys 220
DB 646 CCAACGCTTAAGCTGATGCGCTGAGACCACTGCAAAAGTCTATCTTCTGATGAGAAAG 705
QY 221 ProGlyProHisLysIleGlyGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
DB 706 CCTGCTTCACATTAAGATTCAGGAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
DB 765 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
QY 261 LeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

```

```

DB 825 CTGGATTCGAGAGAAAGATTCCTATGAGGAAATTCATCTGATGCTGCTGCTGCTG 884
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaVal 297
DB 885 ATTAAAGTCGCTAAGCGCTGTAATAATCTGGAGAGCTCATGTTGTTATT 935

```

## RESULT 13

## LOCUS

CG848066 934 bp DNA linear GSS 18-NOV-2003

ZMBBDB0319111.r ZMBBDB Zea mays subsp. mays genomic clone

ZMBBDB0319111.3', genomic survey sequence.

CG848066

CG848066.1 GI:38374927

GSS.

Zea mays subsp. mays (maize)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 934)

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.

and Wing, R.

Sequencing of the maize genome

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: T7

REVERSE: M13r

Place: 0319 row: 1 column: 11

Seq primer: M13r

Class: BAC ends.

Location/Qualifiers

1..934

/organism="Zea mays subsp. mays"

/mol\_type="genomic DNA"

/cultivar="B73"

/sub\_species="mays"

/db\_xref="taxon:4578"

/clone="ZMBBDB0319111"

/lab\_host="DH10B"

/note="Vector: pBel0Bcl1; Site\_1: HindIII; Site\_2:

HindIII; Zea mays L. ssp. mays"

## ORIGIN

## Alignment Scores:

Prod. No.:	1 08e-128	Length:	934
Score:	1285.00 <td>Matches:</td> <td>247 </td>	Matches:	247
Percent Similarity:	95.12% <td>Conservative:</td> <td>25 </td>	Conservative:	25
Best Local Similarity:	86.06% <td>Mismatches:</td> <td>14 </td>	Mismatches:	14
Query Match:	79.17% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	29	Gaps:	0

US-09-931-457a-31 (1-325) x CG848066 (1-934)

```

QY 8 TleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27
DB 72 ATTGCAAGAGATGCTCATATGATTCGCAAAAGCCACTGATATTCGAACAGCTT 131
QY 28 AlaAspGlyCysValAlaAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSer 47
DB 132 GTTGAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
QY 48 ValLysAspArgIleGlyLysSerMetIleAlaAspAlaGluGlyLysLeuIleThr 67

```

Accession	Gene	Species	Length (bp)	Source
D8	ProGly1ySerserValleuilegluProThSerserGlyValenThnglyilegluValaPhe	<i>ProGly1ySerserValleuilegluProThSerserGlyValenThnglyilegluValaPhe</i>	87	192
D5	CCGGGTAAGATGTCCTTAATCGACCAACAGTGTATATACCGGATATGGATTGACATTC	<i>CCGGGTAAGATGTCCTTAATCGACCAACAGTGTATATACCGGATATGGATTGACATTC</i>	311	252
D5	MetAlaAlaAlaATGAGTGTATGlyValSleuilelleThMetProAlaSerMetSerLeuGlu	<i>MetAlaAlaAlaATGAGTGTATGlyValSleuilelleThMetProAlaSerMetSerLeuGlu</i>	107	88
D5	ATGCGACGTCCTTAAGGGTTACAACTCGTATTAAGATGCGCTTCTTCAATGAGTTAGAA	<i>ATGCGACGTCCTTAAGGGTTACAACTCGTATTAAGATGCGCTTCTTCAATGAGTTAGAA</i>	371	312
D5	ArgArgilelleleuValaPheGlyValaGluValleuThrasPProAlaAlaGlyGly	<i>ArgArgilelleleuValaPheGlyValaGluValleuThrasPProAlaAlaGlyGly</i>	127	108
D5	AGGAGATGTCCTCCCGAGCTTTGGAGCTTAACCTGTTCTTCAAGATCCAGCTCGCGC	<i>AGGAGATGTCCTCCCGAGCTTTGGAGCTTAACCTGTTCTTCAAGATCCAGCTCGCGC</i>	431	372
D5	MetValSgIValaValGlnValaGluValleuThrasPProAlaAlaAlaValle	<i>MetValSgIValaValGlnValaGluValleuThrasPProAlaAlaAlaValle</i>	147	128
D5	ATGAAAGGGGCTGTCCTTAAGAGCTGAGAAATTCATGCAAGAACCCCAATTCCTTCA	<i>ATGAAAGGGGCTGTCCTTAAGAGCTGAGAAATTCATGCAAGAACCCCAATTCCTTCA</i>	491	432
D5	LeuGlnGlnPheGluValaPheAlaAlaPheValaPheValaPheValaPheValaPhe	<i>LeuGlnGlnPheGluValaPheAlaAlaPheValaPheValaPheValaPheValaPhe</i>	167	148
D5	CTTACGACATTTGAAACCCCTGCCACCCCAAGATCCATTCACAGCAACCGGACAGAG	<i>CTTACGACATTTGAAACCCCTGCCACCCCAAGATCCATTCACAGCAACCGGACAGAG</i>	551	492
D5	lleTrpIySgIySerserPheGlyValleuSgIySgIySgIySgIySgIySgIySgIy	<i>lleTrpIySgIySerserPheGlyValleuSgIySgIySgIySgIySgIySgIySgIy</i>	187	168
D5	ATATGGAAGGACACAGGGGAGGATATATTAATCTGCTGCTGAGATTTGGACCTGGAGT	<i>ATATGGAAGGACACAGGGGAGGATATATTAATCTGCTGCTGAGATTTGGACCTGGAGT</i>	611	552
D5	ThreThrGlyValaGlyValleuSgIySgIySgIySgIySgIySgIySgIySgIySgIy	<i>ThreThrGlyValaGlyValleuSgIySgIySgIySgIySgIySgIySgIySgIySgIy</i>	207	188
D5	ACAGTACAGAGACAGAGAGATATCTCAAGAGAGAGATCCCAATTAAGCTATGCT	<i>ACAGTACAGAGACAGAGAGATATCTCAAGAGAGAGATCCCAATTAAGCTATGCT</i>	671	612
D5	ValGlnPheValGlnSerProValleuSgIySgIySgIySgIySgIySgIySgIySgIy	<i>ValGlnPheValGlnSerProValleuSgIySgIySgIySgIySgIySgIySgIySgIy</i>	227	208
D5	GTGGAACCCAGTGTGAAGAGCGAGTCTTCTGCTGCTGAGAAACCCGCTGCCATTAAGTAA	<i>GTGGAACCCAGTGTGAAGAGCGAGTCTTCTGCTGCTGAGAAACCCGCTGCCATTAAGTAA</i>	731	672
D5	GlyIleGlyValaGlyPheIleProGlyValleuGluValaPheValaPheValaPheVala	<i>GlyIleGlyValaGlyPheIleProGlyValleuGluValaPheValaPheValaPheVala</i>	247	228
D5	GGACCTCGGTCGCGGATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	<i>GGACCTCGGTCGCGGATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT</i>	791	732
D5	GlnIleSerserPheGlyValleuGluValaPheValaPheValaPheValaPheValaPhe	<i>GlnIleSerserPheGlyValleuGluValaPheValaPheValaPheValaPheValaPhe</i>	267	248
D5	CAATATCAAGAGAGAGAGATCTGCAAACTGCAAGCAAACTGCTTGAAGAAGAGATTA	<i>CAATATCAAGAGAGAGAGATCTGCAAACTGCAAGCAAACTGCTTGAAGAAGAGATTA</i>	851	792
D5	PheValGlyIleSerserGlyValleuGluValaPheValaPheValaPheValaPheVala	<i>PheValGlyIleSerserGlyValleuGluValaPheValaPheValaPheValaPheVala</i>	287	268
D5	TTAGTGTGTAATCAATCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	<i>TTAGTGTGTAATCAATCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT</i>	911	852
D5	GluAlaAlaGlyValleuIle 294	<i>GluAlaAlaGlyValleuIle 294</i>		288
D5	GAATATCTGCGGAGCTCATC 932	<i>GAATATCTGCGGAGCTCATC 932</i>		912
RESULT 14				
LOCUS	AY112530	1258 bp	mRNA	linear
DEFINITION	Zea mays CL5251_1 mRNA sequence.			
ACCESSION	AY112530			
VERSION	AY112530.1			
KEYWORDS	HTC.			
SOURCE	Zea mays			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.			
REFERENCE	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgan, M. and Tingey, S.V. Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes			
TITLE	Unpublished (2002)			
JOURNAL	2 (bases 1 to 1258)			
REFERENCE	Coe, E.H.			
AUTHORS	Direct Submission			
TITLE	Submitted (25-Apr-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			

**COMMENT**

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at [MSL.maizegenetics.org/ZmDB](http://MSL.maizegenetics.org/ZmDB), [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu), [TIGR.zmdb.iastate.edu](http://TIGR.zmdb.iastate.edu), [www.fhg.org/ncbi](http://www.fhg.org/ncbi), [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES	Location/Qualities
source	1. .1258

```

/organisms="Zea mays"
/mol_type="mRNA"
/db_xref="MaidDB:632306"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
library"

```

/note="This sequence is part of a project of 531 assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overo addressing of BACs in conjunction with the Maize Mapping Project"

**ORIGIN**

**Alignment Scores:**

Pred. No.:	9,67e-127	Length:	1255
Score:	1269.00	Matches:	241
Percent Similarity:	88.75%	Conservative:	43
Best Local Similarity:	75.31%	Mismatches:	36
Query Match:	78.13%	Indels:	0
DB:	11	Gaps:	0

US-09-931-457A-31 (1-325) X AY112530 (1-1258)

6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25

Db 1 ACCACATCGCCACGAGTCA<sup>CC</sup>AGCTCATCGCAACACCAATGGTATCTCAAC 60

26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45

Db 61 AACGTGTCAGGGCTCTGTGGCCAATGTCGCTAAGCTCGAGATTATGGAGCCCTGC 127

46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeu 65

Db 121 TGTAGCGTCAAGGACAGGATAGGCTACAGTATGATAACGACGCTGAACAGAGGCTTG 1E

66 IleThrProGlyValIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85

Db 181 ATTACTCCTGGAAGAGTCTTTTGGTGAAGCAACAAGTGGAAACACAGGCATTGGTCTT 24

86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 10

Db 241 GCTTTATGCTGCTTCCAAAGGATATAAGCTGATCTAACATGCTTCATCATGAGC 30

On 106 IeuGluaRArTlleTleuIeuAlapheG]VA]agIuIeuValIeuThraSpProAla 12

201 ATGAGAGAGAGCTCTCTCAGACCTTTGGTCCGAACTTGTCTTACTGATGCTGCA 30

```

136 1xSG1vMet1xSG1vA1aVa1G1n1xSA1aG1uG1u1e1eua1a1b1vSThrProAsnA1a 14

```

361 AAGCGATTGAAGCGGCTTAAAGTAAAGCTACAGAGATTTTAAATTAAGACACCCAATTCT 43

146 Turt[elug]ng]ppheg]u]asnproa]aasnpro]vsval]hi]stvr]glu]thr]thrg]ly 14

421 TATATGCTTACCACTTGATACTGCAACCTAAAGTACATTATGAGACTACTGGT 41

[illegible]

481 CCGGCTCTCCGACCGATTCCTGCTGCAATGGACA 54

108 [1] Atharvashikshaṅgīyasaṁhitā, Nāśanprasaṅgaśloka 20

[illegible]

QY 206 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 225  
 DB 601 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 660  
 QY 226 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 245  
 DB 661 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 720  
 QY 246 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 265  
 DB 721 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 780  
 QY 266 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 285  
 DB 781 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 840  
 QY 286 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 305  
 DB 841 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 900  
 QY 306 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 325  
 DB 901 |||G|V|A|G|U|P|V|A|G|U|S|E|P|P|V|A|L|U|S|E|S|G|Y|L|Y|S|P|P|G|I|P|H|O|I|S|Y|S 960

RESULT 15  
 CK265667 886 bp mRNA linear EST 12-DEC-2003  
 LOCUS EST11745 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 DEFINITION  
 CK265667  
 VERSION  
 CK265667.1 GI:39822645  
 SOURCE  
 EST  
 ORGANISM  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 Buell, C.R., Hart, A., Ziemann, V., Karanayeva, S.A. and Baker, B.  
 (bases 1 to 886)  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other hosts: EST11746  
 Contact: Rodin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
 Source  
 1..886  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAC011"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-TonA"  
 /clone\_1b="potato abiotic stress cDNA library"  
 /note="Vector: pCWSport6.1; Site\_1: EcoRI; Site\_2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.15e-126 Length: 886  
 Score: 1260.00 Matches: 244  
 Percent Similarity: 94.29% Conservative: 20  
 Best Local Similarity: 87.14% Mismatches: 16  
 Query Match: 77.63% Indels: 0  
 DB: 14 Gaps: 0

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

US-09-931-457a-31 (1-325) x CK265667 (1-886)  
 QY 1 MetAlaValAGUArgSerGlyIleAlaIysAspValThrGluLeuIleGlyLysThrPro 20  
 DB 42 ATGGCGGGGAAAGAAATTTGGATTCCCAAGATGTAACTGATTGTTAACTCTCT 101  
 QY 21 LeuValIYrLeuAsnLeuLeuAlaAspGlyCysValAlaIahgValAlaIahgLeu 40  
 DB 102 TTGCTATACCTGAAATTAATGATGATGGGTGTGTGACCGTGTGCGCAAGCTGAA 161  
 QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyIYrSerMetIleAlaAspAla 60  
 DB 162 AGCATGGAGCATGCTCTAGCTTTAAGCATAGCATGCGTTATGATATTAACATGCT 221  
 QY 61 GluIuYsgIYleuIleThrProGlyLysSerValIleuIleGluProThrSerGlyAsn 80  
 DB 222 GAGGAGAAAGGCTCATAAACCCGCGAGAGTGTCTTCATCACTGAACCTTACGATGAAAC 281  
 QY 81 ThrGlyIleGlyLeuLeuAlaPheSerValAlaAlaArgIYrYrLeuLeuIleThrMet 100  
 DB 282 ACTGGGTGAGATGATGCTTCAAGCTCTGCTTAAGGCTCAACCTCATCTATTCAGATG 341  
 QY 101 ProAlaSerMetSerLeuGluArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120  
 DB 342 CTTCTCAATAGCTTGAAGAGAAATTAATTCGCGCTTGTGCGTCAAGTGTGTG 401  
 QY 121 LeuThrAspProAlaIYsGlyMetIYsGlyAlaValGluValaGluGluIleLeuAla 140  
 DB 402 CTTACCGATCCGCAAAAGGATGAAAGGTGCTATTTCAAAGGCTGAAGAGATTAAGGCC 461  
 QY 141 IYsThrProAsnAlaYrIleLeuGluInpHeGluAsnProAlaAsnProIYsValHis 160  
 DB 462 AAAACACCAACTCTCATATTTCTTCAAGCAATTTGAAACCCGCTCAACCAAGATACAC 521  
 QY 161 YrGluThrThrGlyProGluIleThrIYsGlySerAspGlyIYsAlaAspAlaPheVal 180  
 DB 522 TATGAGACCACTGGTCCGATTCGAAAGGCTCAAAAGGATGATCTTGTGAGG 581  
 QY 181 SerGlyIleGlyThrGlyIYrThrIleThrGlyAlaGlyIYsYrLeuIYsGluIuAsn 200  
 DB 582 TCTGATTTGGAACAGAGGACGCAAGTATGTTCAAGCAATTAATTGAGAGAGACAAAC 641  
 QY 201 ProAsnIleYsLeuIleGlyValaGluProValaGluSerProValLeuSerGlyIYs 220  
 DB 642 CCAACGTTTAACTTATAGCCGTGAAACCAAGTGAAGCTATCTTCTGTGTGAGG 701  
 QY 221 ProGlyProIleYsIleGluIYleGlyAlaGlyPheIleProGlyValaLeuGluVal 240  
 DB 702 CTTGCTCAATTAATTCAGGAGTTGCTGTGTTTCAATCCAGGTTTGGAGTT 761  
 QY 241 AsnLeuLeuAspGluValaValaGluIleSerSerAspGluAlaIleGluThrAlaYsLeu 260  
 DB 762 AACCTTATGATGATGTAGTTCAGGTTTCAAGTGAAGATCCATGAATGGCTAAGCTT 821  
 QY 261 LeuAlaLeuYsgIuGlyLeuPheValaGlyIleSerSerGlyAlaAlaAlaAlaAla 280

Mon May 24 08:18:52 2004

us-09-931-457a-31.p2n.rst

Page 15

Db : 822 CTGGCATTGAAAGAGAGATTGCTAGTGGAAATTCATCTGCTGCTCTGCGCGCGCA 881

Search completed: May 22, 2004, 00:35:05  
Job time : 3109 secs

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rn1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 21, 2004, 22:03:57 ; Search time 114 Seconds  
(without alignments)

1582.097 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVERSGIAGKVTBLIGKTP.....LSVLFESVREASMTPEP 325

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-O=/cgm2.1/USPTO/spool/US09931457/rn1at 18052004 121729 20321/app.query.fasta\_1.519  
-DB=Issued Patents NA -QWMT=fast5d -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEXT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931457@cgm2.1 1.56 @rn1at 18052004 121729 20321 -NCPD=6 -ICPD=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgm2.6/p2nmodel/2/ina/5A.COMB.seg:\*  
2: /cgm2.6/p2nmodel/2/ina/5B.COMB.seg:\*  
3: /cgm2.6/p2nmodel/2/ina/6A.COMB.seg:\*  
4: /cgm2.6/p2nmodel/2/ina/6B.COMB.seg:\*  
5: /cgm2.6/p2nmodel/2/ina/PCUTS.COMB.seg:\*  
6: /cgm2.6/p2nmodel/2/ina/backfiles1.seg:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	1179	72.6	975	4	US-09-905-290A-3
2	949.5	58.5	4403765	3	US-09-103-840A-2
3	949.5	58.5	4411529	3	US-09-103-840A-1
4	917	56.5	100848	4	US-09-596-002-39
5	882	54.3	1008	4	US-09-543-681A-3561
6	865	53.3	1008	4	US-09-328-352-1324
7	856	52.7	2247	4	US-09-252-991A-14920
8	856	52.7	2247	4	US-09-252-991A-14450
9	853.5	52.6	981	4	US-09-489-039A-1995
10	816.5	50.3	1830121	4	US-09-557-884-1
11	816.5	50.3	1830121	4	US-09-643-980A-1
12	815	50.2	6115	4	US-09-956-171B-148

13	813	50.1	927	4	US-09-107-532A-271
14	801	49.4	960	4	US-09-134-001C-493
15	777	47.9	10993	4	US-08-961-527-15
16	765.5	47.2	640681	4	US-09-790-988-1
17	761	46.9	927	4	US-09-724-623-10
18	761	46.9	8885	4	US-09-634-238-26
19	650.5	40.1	912	4	US-09-543-681A-3548
20	628.5	38.7	921	4	US-09-489-039A-2004
21	601	37.0	66986	4	US-09-596-002-39
22	598.5	36.9	1149	4	US-09-252-991A-12995
23	598.5	36.9	3486	4	US-09-252-991A-12477
24	594	36.6	924	4	US-09-328-352-4053
25	593	36.5	933	4	US-09-540-236-1422
26	586	36.1	36800	4	US-08-311-731A-139
27	552.5	34.0	927	4	US-09-107-532A-2975
28	550	33.9	2500	3	US-09-318-448-14
29	550	33.9	2542	1	US-08-120-960-1
30	550	33.9	2542	4	US-09-347-878-8
31	531.5	32.7	1524	4	US-10-012-762-14
32	531.5	32.7	1524	4	US-09-704-036B-14
33	530	32.0	840	4	US-09-134-000C-1996
34	487.5	30.0	8536	4	US-08-956-171B-278
35	485	29.9	1377	4	US-09-489-039A-2013
36	481	29.6	1413	4	US-09-252-991A-437
37	481	29.6	1437	4	US-09-252-991A-364
38	479.5	29.5	939	4	US-09-134-000C-121
39	474	29.2	939	4	US-09-134-000C-1083
40	463	28.5	741	4	US-09-252-991A-13320
41	451.5	27.8	936	4	US-09-252-991A-8456
42	446.5	27.5	495	4	US-09-252-991A-15179
43	408	25.1	560	4	US-09-621-976-519
44	404	24.9	536165	4	US-09-214-808-1
45	396	24.4	600	4	US-09-252-991A-13162

## ALIGNMENTS

RESULT 1  
US-09-905-290A-3  
Sequence 3, Application US/09905290A  
Patent No. 6605459  
GENERAL INFORMATION:  
APPLICANT: Rice, John  
APPLICANT: Crawford, Beth  
APPLICANT: Nye, Gordon  
TITLE OF INVENTION: METHODS FOR MEASURING CYSTEINE AND DETERMINING CYSTEINE SYNTHASE  
FILE REFERENCE: Docket NO. 660545920900US  
CURRENT APPLICATION NUMBER: US/09/905.290A  
CURRENT FILING DATE: 2001-07-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 975  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURES:  
NAME/KEY: misc feature  
OTHER INFORMATION: AJ011603 GenBank  
US-09-905-290A-3  
Alignment Scores:  
Pred. No.: 7.78e-130  
Score: 1179.00  
Percent Similarity: 84.88%  
Best Local Similarity: 68.21%  
Query Match: 72.64%  
Gaps: 0  
US-09-931-457A-31 (1-325) x US-09-905-290A-3 (1-975)  
1 MetalaValGuaGserGlyTlaAlaIysaPvaThrgIueuIlleglyLysThrPro 20  
|||||

Mon May 24 08:18:51 2004

US-09-931-457a-31.p2n.rni

Page 2

Db	1	ATGAGAGAGGACCGGGTGTTCGATCCAGAGATGATGCAACTCAATTGTATAGGTAAACCCCA	60
Qy	21	LeuVal1YrLeuAsnIyIleuAlaAspGlyCYsValAlaArgValAlaAlaIyIleuGlu	40
Db	61	ATGGTATATATCTGAACAACAATTGGATATGGTGTGGAGCTCGATGATGACGTAAAGCTTGAA	120
Qy	41	LeuMetGluProCysSerSerValIyAspArgIleGlyTyrSerMetIleAlaAspAla	60
Db	121	ATGATGAGACCTTGCTGCTACTGATCCAGAGCCAAATTGGCTATAGCTATATTAAGATGCA	180
Qy	61	GluGluIySGlyLeuIleThrProGlyIyVSerValIleuIleGluProThrSerGlyAsn	80
Db	181	GAGAGACAAAGGATTTGATTACTCTCTGGGAAAGAGACACTATATTAAGAGCTACCTCTGGTAA	240
Qy	81	ThrGlyIleGlyLeuAlaPheMetAlaAlaIleArgGlyTyrIyIleuIleIleThrMet	100
Db	241	ACCGGAGTTGGTTTAGCTTCACTGATCCGGTGCAGCTTAAGGTTTCAAAAGTGTCTCAAAATG	300
Qy	101	ProAlaSerSerSerLeuGluIyArgArgIleIleLeuLeuAlaPheGlyValAGluLeuVal	120
Db	301	CCCTCAATCAATGAGCCTTGAGAGAAAAAATCATTTCTTTAGCATTTAGTGGCGGAGGTCTAC	360
Qy	121	IeuThrAspProAlaIySGlyMetIySGlyValAlaGluValAlaAGluGluIleLeuAla	140
Db	361	CTCAAGATCTCTGATGAAGGCGCTTCAAGAAATATGACAAAGCTGAGAAAGATATATGATAC	420
Qy	141	IyThrProAsnAlaTyrIleLeuGluIyPheGluAsnProAlaAsnProIyValHis	160
Db	421	AAAAATCCAAATGATGATCATGCTAGAGAAAGTTCAAAAATCTTCAAAACCGCAAACTCAT	480
Qy	161	TyrGluThrThrGlyProGluIleTyrPlySGlySerAspGlyIyAlaAspAlaPheVal	180
Db	481	TATGAGAACCGGGGTCCAGGATATATGAGAGACTCTCCAGGGGAAAGTGAACATTTGGTT	540
Qy	181	SerGlyIleGlyIyThrGlyGlyThrIleThrGlyAlaGlyIySerIyLeuIySGluGluAsn	200
Db	541	GCGGATTTTGAATCGGTGGAGACGCTTCCGAGATCCAGAAAGATTCCTCAAGGAGAGAT	600
Qy	201	ProAsnIleIyIleuIleGlyValAGluProValGluSerProValLeuSerGlyIyAs	220
Db	601	AAAGACTTTTAAAGTTTATGCTGTGGAACTTACAGAAAGTCCGTTATATAGTGGAGGCAA	660
Qy	221	ProGlyIProHisIyIleGluIyIleGlyIyAlaGlyPheIleProGlyValIleuGluVal	240
Db	661	CCGGGTACACATTTGATCCAAAGGTATTTGGGGCTGCACTATCCCAACAATTTGGATATTC	720
Qy	241	AsnIleuAsnAspGluValIyGluIleSerSerAspGluAlaIleGluThrAlaIyLeu	260
Db	721	AACTTTCTTATATAGTATGCATCAAGTGAACAAAGTGGAAAGCAATTGAAACAGCCAAACTT	780
Qy	261	LeuAlaLeuIySGluGlyIyLeuPheValAGlyIleSerSerGlyIyAlaAlaAlaAlaAla	280
Db	781	CTTGCCCTGAAGAAGGATTACTGGTGGGAAATATCTTCTGGAGCTGCTGAGCGCGTGGC	840
Qy	281	PheGluIleAlaIyAspArgProGluAsnAlaGlyIyIleuIleValAlaValPheProSer	300
Db	841	ATTAAGGTGGCAAGGCGCCAGAAACGCGGCAATCTATAGTGTGATTTCCCTGAC	900
Qy	301	PheGlyGluArgTyrLeuSerSerValIleuPheGluSerValArgArgGluIleGluSer	320
Db	901	GGTGGGGAAGCTTAACCTATGACCTCATTTGTCGAATCAGTCAAGACATGAAGCAGAGAT	960
Qy	321	MetThrPheGlu 324	
Db	961	TTTCGAATTCAA 972	

```

1  APPLICANT: FRASER, Claire M.
2  APPLICANT: VENTER, John C.
3  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
4  TITLE OF INVENTION: TUBERCULOSIS
5  FILE REFERENCE: 24366-20007.00
6  CURRENT APPLICATION NUMBER: US/09/103,840A
7  CURRENT FILING DATE: 1998-06-24
8  NUMBER OF SEQ ID NOS: 2
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 2
11 LENGTH: 4403765
12 TYPE: DNA
13 ORGANISM: Mycobacterium tuberculosis
14 FEATURE:
15 OTHER INFORMATION: CDC 1551
16 OTHER INFORMATION: "n" bases at various positions throughout the sequence
17 OTHER INFORMATION: represent a, t, c or g
18 US-09-103-840A-2

```

```

Alignment Scores:
Pred. No.:      7.71e-97      length: 4403765
Score:          949.50        Matches: 181
Percent Similarity: 76.95%    Conservative: 56
Best local Similarity: 58.77%  Mismatches: 68
Query Match:    58.50%        Indels: 3
DB:             3            Gaps: 3
US-09-931-457A-31 (1-325) x US-09-103-840A-2 (14403765)

```

QY	8	lalelalyssapvaltthngulneullellylthsrprouvaltytleasnytleu	27
Db	2604610	ATGCCGAGACATCACTCCACTATGCGGCGACACCCGTCGTCCATCCGCCGCACTC	2604665
QY	28	AlaaspGlyCyValAlaIarGValAlaAlaLysLeuGluLeuMetGluProCysSer	47
Db	2604670	ACCGACGGCCCGCTTCCGACATCGTCGCCAAGCTGGAAATTCTTCAACCCGGCCACAGC	2604729
QY	48	VallyasprgrtleglytyrserMetIleAlaaspIaGluGluysGlyLeuIlethr	67
Db	2604730	GTAATAAACCTATATCGGCGTTCGCACTCTCCAGCGGCCAGACGAGGAGTTTATCAAG	2604789
QY	68	ProGlyLysSerValLeuIleGluProThrSerGlyasnThrGlyIleGlyLeuAlaIphe	87
Db	2604790	CCG---GACACGATCATCTTCGAAACCAACGAGGTAAACCCGACATCGCCTTGCCATG	2604846
QY	88	MetalAlaIaIaIarGlytyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu	107
Db	2604847	GTTTCGCGGACGACGCGGCTACCGGTGCGTGCATGCCATGCCCCGAGCATGAGTCTGGAG	2604906
QY	108	ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaIaLysGly	127
Db	2604907	CGCGGATTTGTCTGCGCGCATACCGGCTGAGTCACTCACTCCGGTCCGACAGGC	2604966
QY	128	MetLysGlyAlaIaValGluGlyAlaGluGluIleLeuAlaLysThrProAsnAlaTyrlle	147
Db	2604967	ATGTCAGGTGCTCATTCGCCAAGCTGTAGGAG---CTGGCCCAAGACCGATCAACGCTACTTC	2605023
QY	148	Leu---GlnGlnPheGluAsnProAlaAspProLysValHisGlytyrGluThrThGlyPro	166
Db	2605024	GTCGCCAGACATTCGAGAACCCCGCGCAACCCGCCATCATCGCGTCAACGACCGCGGAG	2605083
QY	167	GluIleIleTrpLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly	186
Db	2605084	GAGGTCTGGGTACACACCGACGCGAAGTGCACATCTGTCGTCCGGGAGTGGGACCGGT	2605143
QY	187	GlyThrIleThrGlyAlaGlyLysTyrlLeuLysGluGlnAsnProAsnIleLysLeuIle	206
Db	2605144	GGCACCATTCACCGGCGTCCGCGAGTTCATCAAGAAACGCAACCGTCCGCGCCGGTTCGTG	2605203
QY	207	GlyValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIle	226
Db	2605204	GCCGTAAAGCCGCGCGCGTCCGCGTCTTCTGTGGTGGCAAGAAAGGACCGACCCGATC	2605263



QY 227 GINGLYIIEGLYALAGLYPHEIIEPROGLYVALLEUHLVALASNULEUENUSPGLUVAL 246  
DB 2605264 CAGGCGATCGCGCGCGGTTCTGCTCCCGGCTACTGACCAAGACTTATGCAAGATC 2605323  
QY 247 VALGNIIESESESPGLUVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 266  
DB 2605324 ATTACCGTCGATACCAAGACCGGCTCAACGTCGCGCGCGCTGCGCGGACGAGGCGC 2605383  
QY 267 LEUPHEVALGYLIESESESPGLYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 286  
DB 2605384 TTGCTGTCGCGATCTCTCGGCGCGGCTCAACGTCGCGCTTCAAGTGGCGCGCGC 2605443  
QY 287 PROGLUASNALAGLYLEUENLEVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 306  
DB 2605444 CAGGAGAACCGCGGAGGCTATCGTCGAGTCTCTCCGACTTCGCGGACGATATCTG 2605503  
QY 307 SERSEVALLEUPHEGLUSERVAL 314  
DB 2605504 AGCACACCGTGTTCGCGGACGTC 2605527  
RESULT 3  
US-09-103-840A-1  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 24366-20007.00  
CURRENT FILING DATE: US/09/103,840A  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1  
Alignment Scores:  
Pred. No.: 7.74e-97 Length: 4411529  
Score: 949.50 Matches: 181  
Percent Similarity: 76.95% Conservative: 56  
Best Local Similarity: 58.77% Mismatches: 68  
Query Match: 58.50% Indels: 3  
DB: 3 Gaps: 3  
US-09-931-457a-31 (1-325) x US-09-103-840A-1 (1-4411529)  
QY 8 IIEALYASPAVALTHRGULENIIIEGLYLYSETHPROLEUVALYRIEUNUSL 27  
DB 2608800 ATCGCGAGAGACATCAACCACTCATCGGCGCAACCGCTGCTCCGACTGCCCAAGATC 2608859  
QY 28 ALASPGLYCYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 47  
DB 2608860 ACCGAGCGCGCGCTTCCGACATCGTCGCGAGCTGGAATCTTCAACCGCGCAACGCG 2608919  
QY 48 VALIYASPAVALIIEGLYLYSETHPROLEUVALYRIEUNUSL 67  
DB 2608920 GTAAAGACCGATCGGCGTTCGACATCGTCGAACGCGCGGACGAGGATTCGATCAAG 2608979  
QY 68 PROGLYLYSESESPGLUVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 87  
DB 2608980 CCG--GACGAGATCTTCGAAACCGAGCGGTAACCGCGCATCGCGCTGCGCGAG 2609036  
QY 88 METALAALAIAGLYLYLYBLEULIEIETHMEPROALASERSETSELEUGLU 107  
DB 2609037 GTTGGCGCGGACGCGGCTACCGGTGCTGCTGACCATTCGCGAGACGATGACTGAG 2609096

QY 108 ARGATRIIELELEUENALAPHEGLYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 127  
DB 2609097 CCGCGAGTGTGCTCGCGGATACGTCCTGATCACTCATCTCACTCCGCGTCCGAGCGC 2609156  
QY 128 NEULYSGLYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 147  
DB 2609157 ATGTCAAGTGTGCTCGCGGATACGTCCTGATCACTCATCTCACTCCGCGTCCGAGCGC 2609213  
QY 148 LEU---GINGLIIESESESPGLUVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 166  
DB 2609214 GTGCGCGAGATTCGAAACCGCGGACCGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGG 2609273  
QY 167 GIULETTPLYSESESPGLYLYSILEAPALAPHEVALSERGLYIIEGLYLYTHRG 186  
DB 2609274 GAGCTCTGCGGATCAACGAGCGGACGAGTTCGATCTGCGCGGATTCGCGGATTCGCGGATTCGCGG 2609333  
QY 187 GIUTRIIELELEUENALAPHEGLYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 206  
DB 2609334 GCGACATCAACCGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGG 2609393  
QY 207 GIYVALIIEGLIUTRHALALYSESESPGLUVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 226  
DB 2609394 GCGGTAGAGCGCGCGCGCTGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGGATTCGCGG 2609453  
QY 227 GINGLYIIEGLYALAGLYPHEIIEPROGLYVALLEUHLVALASNULEUENUSPGLUVAL 246  
DB 2609454 CAGGCGATCGCGCGGTTCTGCTCCCGGCTACTGACCAAGACTTATGCAAGATC 2609513  
QY 247 VALGNIIESESESPGLUVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 266  
DB 2609514 ATTACCGTCGATACCAAGACCGGCTCAACGTCGCGCGCTGCGCGGACGAGGCGC 2609573  
QY 267 LEUPHEVALGYLIESESESPGLYVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 286  
DB 2609574 TTGCTGTCGCGATCTCTCGGCGCGGCTCAACGTCGCGCTTCAAGTGGCGCGCGC 2609633  
QY 287 PROGLUASNALAGLYLEUENLEVALIIEGLIUTRHALALYSELEUENALALEUYSGLUGLY 306  
DB 2609634 CAGGAGAACCGCGGAGGCTATCGTCGAGTCTCTCCGACTTCGCGGACGATATCTG 2609693  
QY 307 SERSEVALLEUPHEGLUSERVAL 314  
DB 2609694 AGCACACCGTGTTCGCGGACGTC 2609717  
RESULT 4  
US-09-596-002-39  
Sequence 39, Application US/09596002  
Patent No. 6632636  
GENERAL INFORMATION:  
APPLICANT: Lagace, Robert, E.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
FILE REFERENCE: PM-0008-4 US  
CURRENT FILING DATE: US/09/596,002  
PRIOR APPLICATION NUMBER: 60/140,121  
PRIORITY FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PERL Program  
SEQ ID NO 39  
LENGTH: 100848  
TYPE: DNA  
ORGANISM: M. catarrhalis  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte template ID No. 6632636 39  
PUBLICATION INFORMATION:  
US-09-596-002-39  
Alignment Scores:  
Pred. No.: 1.29e-95 Length: 100848  
Score: 917.00 Matches: 182





Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rn1

**Page 6**

QY 26 LysLeuLAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45  
 Db 1336 CGCATGCGCCCGCGCGCGGTACCC--ATCTCGGCAAGATCGAAGGCGCAACCCGCG 1392  
 QY 46 SerSerValIlyAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluValGlyLeu 65  
 Db 1393 TATTCGGTGAAGTGCCTGGATCGCGCGCAATGATCTGGAGATGCGGAAGCGCGGAGA 1452  
 QY 66 IleThrProGlyIlySerSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85  
 Db 1453 CTGAAGTCCTGGCAAGAAC--CTGGTCAGCGCACTCCCGGACACCCGGCATGAGCTTC 1509  
 QY 86 AlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSer 105  
 Db 1510 GCCTTCGTCGCGCGCGCGCGCTACAAATCATGATCCGACATGCGCGGCTGATGAGC 1569  
 QY 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyIleGluLeuValLeuThrAspProAla 125  
 Db 1570 CTGAAGCAACCAAGATGCTCAAGCGCTGGAGCGGAGGCTGGTCTGACCAACCGGCG 1629  
 QY 126 LysGlyMetLysGlyAlaValaGluLysAlaGluGluIleLeuAlaLysThrProAsnAla 145  
 Db 1630 AAGGGCATGAAGGGTGGCATCGAAGAGCGAGGAATCGTGGCGGGGAATCCCGGGAAG 1689  
 QY 146 TyrIleLeu---GlnGlnPheGluAspProIleAspProCysValHisTyrGluThrThr 164  
 Db 1690 TACTTCATGACCGCAGACAGTCTGACACCCCGGCGAACCAGCGATCGATGAAGAACAACC 1749  
 QY 165 GlyProGluIleTProLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGly 184  
 Db 1750 GGCCCGGAGATCTGAAGCATACCAAGAGGCGCGGAGACGTGCTGATCTCGGAGATCGCG 1809  
 QY 185 ThrGlyGlyThrIleThrGlyAlaGlyLysTyrIleuLysGlu---GlnAsnProAsnIle 203  
 Db 1810 ACCGCGCGACCTCCACCGCGGTCTCGGCTCATCAAGAAACACCCCGGCGAACCAGTAC 1869  
 QY 204 LysLeuIleGlyValaGluProValaGluSerProValLeuSer-----GlyGly--- 219  
 Db 1870 CTGCGCGTGGGAGTGAACCGGTGACCTCGCGCGGTGATCAACGACCCCTGCGCGGAG 1929  
 QY 220 -----LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyVal 237  
 Db 1930 GAGGTCAAGCGCGCGCGCGCGCAAGATCCAGAGGACATCGCGCGCTTCGTGTCGAAGAC 1989  
 QY 238 LeuGluValaAsnLeuLeuAspGluValaValaGlnIleSerSerAspGluAlaIleGluThr 257  
 Db 1990 CTGACCTCTGCTGATGACCGCGGTCCAGAAAGATCGCGCGACGAGCAAGCAAGACATG 2049  
 QY 258 AlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAla 277  
 Db 2050 GCGCTGCGCCCTGAATGCAGAGAAGGCAATCTCTCGCGCATCTCTCTCGCGCGCGCATG 2109  
 QY 278 AlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAla 297  
 Db 2110 GCGCGCGGCGTCCGCTGCGCGAGAGCGCAAGATCGACGGCGAAGCAATGCTGTATC 2169  
 QY 298 PheProSerPheGlyGluArgTyrIleuSerSerValLeuPheGluSerValArgArgGlu 317  
 Db 2170 CTCCCGCATTCGCGGAGCGGCTATCTTTCAGAGCATGCTTTCGAGCGGCTGTTCAGCGAA 2229  
 QY 318 AlaGlu 319  
 Db 2230 CAGGAA 2235

RESULT 8  
 US-09-252-991A-14450/C  
 ; Sequence 14450, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubinfeld et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

[illegible][illegible]

Db 411 CTGACCTGTGCTGTGTGACCCGGGTCGAGAGATGTGGGAGACGACGAGGAAAGACATG 352

Qy 258 AAlaIysLeuLeuAlaLeuIysGluGlyLeuPheValGlyTlIseSerGlyAlaAlaIa 277

Db 351 GCGCGCCGCTGTGATGCGAGAGAGAGGACATCTCTGGGAGATCTCTCGCGCGGGAGATG 292

Qy 278 AAlaIaAlaPheGlnIleAlaIysArgProGluAsnAlaGlyIysLeuIleValAlaVal 297

Db 291 GCGCGCCGCGTGTGCGCTGCGCCGAGAGCCGAGACATGACAGGCGAAGACATGTGTGATC 232

Qy 298 PheProSerPheGlyGluArgTyrIleuSerSerValIleuPheGluSerValArgGlu 317

Db 231 CTCCCCGATTCGCGGAGCGGCTATCTTTCGAGCATGCTCTTCGACGCGCTGTTCAGCGAA 172

Qy 318 AlaGlu 319

Db 171 CAGGAA 166

RESULT 9  
US-09-489-039A-1995  
Sequence 1995, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Bretton et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709, 2004001  
CURRENT FILING DATE: 2000-01-27  
PRIORITY FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 1995  
LENGTH: 981  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-1995

Alignment Scores:  
Pred. No.: 2,54e-91 Length: 981  
Score: 853.50 Matches: 175  
Percent Similarity: 71.52% Conservative: 56  
Best Local Similarity: 54.18% Mismatches: 79  
Query Match: 52.59% Indels: 13  
DB: 4 Gaps: 5

US-09-931-457A-31 (1-325) X US-09-489-039A-1995 (1-981)

Qy 6 SerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsn 25

Db 13 AGTAGAGATTTTGAAGACACACTCCCTGACTATCGGCATACACCGCTGTGACACTGAAC 72

Qy 26 LysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCys 45

Db 73 CGCATCGGTAAAGGC-----CCTATCTCGCGCAAGGTAGAACTCGTGAACCCGAGC 123

Qy 46 SerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluIysGlyLeu 65

Db 124 TTTACCGTAAATATGCGGTATCGCGCGGCAATATGATCTGGGATGCTGAAAAAGCCGCGGT 183

Qy 66 IleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeu 85

Db 184 CTGAAGACTCGGCGTGGAA---CTGTGCAAGCCGACCGCAAGCCGATATGCGCTC 240

Qy 86 AlaPheMetAlaAlaAlaArgGlyTyrIleLeuIleIleThrMetProAlaSerMetSer 105

Db 241 GCCTACGTGCGCGGGGCGCGCGCTACAAACATGACCCCTAACCAATGCGGAAACATGAGC 300

Qy 106 LeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAla 125

Db 301 ATCGAGCGCTGTAAAGCTGTGAAAACCCCTGGGCGCAACTGTGTGTGACCGAAGGCGGC 360

```

Oy 126 TysGIwMelYSGIYALValGInLYsAlagInuILeuaLIaLYsThProASnAla 125
Db 361 AAMGGCATGAAGGCCCTATCCAGAAAGCGGAAGAGATTGTCGCACAGAACCCGGAA 420
Oy 146 Tyr---ILeLeuGInLIaPheLIuSnpProLIaSnProlYsValLIhIstYrGLuThrThr 145
Db 421 TTCTCTCTCTGTGAAGATTGTCAGCAACCCGGCTAAACCCGGAATTCACAGAAACATACC 480
Oy 165 GIYProGInLIeTTPheYSGIYSerASpGILYsIleASpAlaPheValISerGIYIleGIY 184
Db 481 GGCCCGGAATCTGGGAAGATACCCGATGCTCAGGTGATGATGTTATCTCCGGGGTAGGT 540
Oy 185 ThrGIYGIYThrILeThnGIYLaGIYstYrLeuLYsGIuGInASnpProASnILeLYs 204
Db 541 ACCGGGGAAACCGTGAACCGCGCTCAGCGCTATATCAAAAACCAACCAAGTAAAAAAGAT 600
Oy 205 LeuILe-----GIYALGIuPProVALGIuSerProVALLeuSerGIYGIY----- 219
Db 601 CTGATACCGCGTCCGCGTTGAACCGAACCGATTCGCCGGTGAATGCCCGACGCGCTGCGGG 660
Oy 220 -----LYeProGIYProHISYsYsIleGInGIYIleGIYAlaGIYpHeLIeProGIY 236
Db 661 GAAGAGCTCAAGCGCGGTCCGACAAATTCAGGGTATCGCGCGCGCTTCATCTCGGC 720
Oy 237 ValLeuGIuValASnLeuASpGILuValValGInLIeSerASpGIuAlaIleGIu 256
Db 721 AACTGATGATCGAAGTGGTGTGCAAAAGTATCGGCATTACCAACGAGAGATATCTCC 780
Oy 257 ThrAlaLYsLeuLeuAlaLeuLIeSInuLYLeuPheValIGIYIleSerASpGIYAlaAla 276
Db 781 ACGCGCGCGCGCTGATGGAAGAAAGACATCTCGCGGGTATCTCTCCGGGGCCGCC 840
Oy 277 ALAlaAlaAlaPheGInLIeAlaLYsArGPProGInuSnaLIaGIYLYsLeuILeValAla 286
Db 841 GTTGCGCGCGCGGTGAAGCTGCAAGAAATGAGCGCTTACCAATTAAGAAATATGTGTT 900
Oy 297 ValPheProSerPheGIYGIuLYrYrLeuSerSerValLeuPheGIuSerValArgArg 316
Db 901 ATCTGCTCTCTCTCGAGGTAGCGTTATTTAAACACCGCGCTGTTGCGCGATCTCTCACT 960
Oy 317 GIuAlaGIu 319
Db 961 GAGAAAGAG 969

RESULT 10
US-09-557-884-1
Sequence 1, Application US/0957884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995

```

ATTORNEY/AGENT INFORMATION:  
 NAME: Michelle S. Marks  
 REGISTRATION NUMBER: 41,971  
 REFERENCE/DOCKET NUMBER: PB186P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-557-884-1

Alignment Scores:  
 Pred. No.: 1,03e-81 Length: 1830121  
 Score: 816.50 Matches: 172  
 Percent Similarity: 70.16% Conservative: 49  
 Best Local Similarity: 54.60% Mismatches: 83  
 Query Match: 50.31% Indels: 11  
 DB: 4 Gaps: 6

US-09-931-457a-31 (1-325) x US-09-557-884-1 (1-1830121)

QY 8 IlaAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 1165136 ATTATATGCAACATTCTTATTCATGCAATACGCCCTTGCTTAAACACTTT 1165195  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47  
 DB 1165196 -----GGCCATAACGGTATGCTGCTGTAATAAATTGAGTGTAAACCACTACAGC 1165249  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
 DB 1165250 GTAAATATGCGCGATTTGGCGCATATGCTGTCAGACAGCAAAAAGATGACAGCTTCA 1165309  
 QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 1165310 AAAGGAAAGAGATT--GTAGATGCAACAGAGTGTAAACGCGCATTTGGCTTAT 1165366  
 QY 88 MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107  
 DB 1165367 GTTCCGCGCGGTAGAGGTATTAATAATCAAGCTCATGTCAGAAACAAATGAGCTTGA 1165426  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyValGluLeuValLeuThrAspProAlaLysGly 127  
 DB 1165427 AGAAAGCGCTTATTTGTCGCGATTGGGTAAATTAGTGTACCGAAGCGCGAAAGCA 1165486  
 QY 128 MetLysGlyAlaValGluLysValGluGluIleLeuAlaLysThrProAlaAlaTyrIle 147  
 DB 1165487 ATGAAGGTGTATTTGCAAGAGCAAGAAATTGTCTCTGATCCAGCGCTATGTC 1165546  
 QY 148 ---LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro 166  
 DB 1165547 ATGCTTAAACATTGAAAAATCCAGCCACCAATTCATCATGAGAAACAAGCGGCT 1165606  
 QY 167 GluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 186  
 DB 1165607 GAAATTTGAAAGATAGCGATGCGAAAGTCATCTGTGTGTCGCGTGAAGAACGTT 1165666  
 QY 187 GlyThrIleThrGlyAlaGlyLysLysLeuLys---GluGlnAsnProAsnIleLysLeu 205  
 DB 1165667 GGTTCGATTACGGGCAATTTCTCGCGCATTAATTAATTTGGTAAACAATACCTTCT 1165726  
 QY 206 IleGlyValGluProValGluLysSerProValLeuSer-----GlyGly----- 219  
 DB 1165727 GTTCCCGCTTAGCGCAGTGAATCTCCGGTCAATTAGTCAAACTTAGCGCGGTGAAGAA 1165786  
 QY 220 LysProGlyLysProHisLysIleGlnGlyTyrIleGlyAlaGlyPheIleProGlyValLeuGlu 239  
 DB 1165787 AACCCAGGCCACACAAATTAATCAAGATACGATGCGGTTCATTTCCAAAAATTATGAT 1165846

QY 240 ValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLys 259  
 DB 1165847 TTATCTATTATTAATGCGGTGAAAACCTGTGATAGTGAACCGCATTCGCCACACTGCT 1165906  
 QY 260 LeuLeuAlaLeuLysGluGlyLysPheValGlyIleSerSerGlyValAlaAlaAlaAla 279  
 DB 1165907 CGCTTAATGCGGAGAAAGAAATCTTTCAGGATTTTCATCTGTCGACGCTGCGCGCT 1165966  
 QY 280 AlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValAlaPhePro 299  
 DB 1165967 GCAGATCGCTTAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1166026  
 QY 300 SerPheGlyLysArgTyrLeuSerSerValLeuPheGluSerVal 314  
 DB 1166027 TACGCTCTGAAACGCTACTTAAGCACACACACTGTTGAAGGATT 1166071

RESULT 11

US-09-643-990A-1

Sequence 1, Application US/09643990A  
 Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of  
 The Haemophilus influenzae Rd Genome, Fragments  
 Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS V6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Alignment Scores:

Pred. No.: 1,03e-81 Length: 1830121

Score: 816.50 Matches: 172

Percent Similarity: 70.16% Conservative: 49

Best Local Similarity: 54.60% Mismatches: 83

Query Match: 50.11% Indels: 11  
DB: 4 Gaps: 6

US-09-931-457a-31 (1-325) x US-09-643-990A-1 (1-1830121)

QY 8 ILeAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
DB 1165136 ATTATGAGACAAATTGTTATTCATCGCAAAATGCGCGCTGTCGCTTAAACACTT 1165195  
QY 28 AlaAspGlyCysValAlaIaIaValAlaAlaLysLeuGluLeuMetGluProCysSer 47  
DB 1165196 -----GGCCATTAACGGTAATGTTGGTGAATAAATGTAAGTCTTAACCCAGCTACAC 1165249  
QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67  
DB 1165250 GTAAATATGCGCTATTTGGCGCAATTTGCTGTCGCAAGCAAGAAATGCAAGCTCACA 1165309  
QY 68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
DB 1165310 AAAGGAAAGAGATT---GTAGATGCAACGAGGTGTAACAGGCGCATTTGCTTAT 1165366  
QY 88 MetAlaAlaIaIaArgGlyTyrTyrLeuLeuIleThrMetProAlaSerMetSerLeuGlu 107  
DB 1165367 GTTGGGGGCTAGAGGTTATTAATACGCTCCTATGCGCAAAACATGATCTTGA 1165426  
QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspProAlaLysGly 127  
DB 1165427 AGAAACCGCTATTATGCGGATGGGTGTAAATTAGTGTACCGCAAGCGCAAAAGGA 1165486  
QY 128 MetLysGlyAlaValGluLysAlaGluGluIleLeuAlaLysThrProAlaIaTyrIle 147  
DB 1165487 ATGAAAGGTCTATGTCGCAAGAGAGAAATGTCCTCTGATCCAGCGCTATGTC 1165546  
QY 148 ---LeuGluGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyPro 166  
DB 1165547 ATGCTTAAACATTTGAAATTCGACCAACCAACATTCACGAAACAAACAGGGCT 1165606  
QY 167 GluIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 186  
DB 1165607 GAAATTCGAAAGATGCGATGCGAAATGTCATGTTGTTGTCGCGGTGGAACAGGT 1165666  
QY 187 GlyThrIleThrGlyAlaGlyLysTyrLeuLys---GluGlnAsnProAlaIleLeu 205  
DB 1165667 GGTTCGATTCAGGCGCATTTCTCCGCCCATTAATTTGATTTTGGTAAACAAATCACTCT 1165726  
QY 206 IleGlyValGluProValGluSerProValLeuSer-----GlyGly----- 219  
DB 1165727 GTTGGCGCTTGAACCGATGGAATCTCCGCTCATAGTCAACCTTACCGCGTGAAGAGTA 1165786  
QY 220 LysProGlyProHisLysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGlu 239  
DB 1165787 AAACCGAGCCCAACAAATTCAGATTCGATCGCGGTTTCATTTCCAAAATTTAGAT 1165846  
QY 240 ValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaLys 259  
DB 1165847 TTATCTATATATGATCGCGTGAACCTGTGATAGTATACCGCACTGCCACAGCTCGT 1165906  
QY 260 LeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAla 279  
DB 1165907 CGCTTAATGCGGAAAGAGAAATTTTCGAGGATTTTATCTGATGACGCTGCGGCT 1165966  
QY 280 AlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPhePro 299  
DB 1165967 GCAAGTCCCTTAGCTAAATTAACCAAGATTCGCTGATAATCACTGTTGATTTTGCT 1166026  
QY 300 SerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerVal 314  
DB 1166027 TCAAGCTCGAAGCTACTTAAACAACACTGTTTGAAGGAT 1166071

GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6115 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-08-956-171E-148  
Alignment Scores:  
Pred. No.: 17e-85 Length: 6115  
Score: 815.00 Matches: 164  
Percent Similarity: 72.94% Conservative: 57  
Best local Similarity: 54.13% Mismatches: 76  
Query Match: 50.22% Indels: 6  
Gaps: 5  
US-09-931-457a-31 (1-325) x US-08-956-171E-148 (1-6115)  
QY 11 AspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeuAlaAspGly 30  
DB 4759 AATATTACTCAAAATTATTTGGCGGTACCGGTATGTCAAATTGAGAAATGATGATGAC 4818  
QY 31 CysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSerValLysAsp 50  
DB 4819 AATGACCAAGTCTTATGTAAATTTGATTAATCAAAATCAAGGTCTCTGTAAGCAT 4878  
QY 51 ArgIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThrProGlyLys 70  
DB 4879 AGAATTCCTTAGCAATGATTAAGAAAGCAAGCGGAGCAAGCAAAATTAACCTGCGCAT 4938  
QY 71 SerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPheMetAlaIa 90  
DB 4939 ACAATT---GTAGAACCAACAAAGTGTAAATACAGTATCGGTTAGCATTTGATGTCCT 4995  
QY 91 AlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGluArgIle 110

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rml

Page 10

4996 GCTAAAGATATTAAGCATATTACTATGCCCCGAAACATGAGCCAGGCGTCTAAT 5055  
111 ILeuLeuAlaPheGlyAlaGluValLeuThrAspProAlaGlySerGlyMetGlyGly 130  
5056 TTAATTAAGGATACCGGTGCGAATTAAGTTTAAACGCTGATGATGAGGATGAAAGGT 5115  
131 AlaValGlnLysAlaGluGluLeuAlaGlyThrProAlaGlyTyrIleLeuGln 149  
5116 GCATTTAAAGGATTAAGATTTG-----AAGAGAAACATGGTTCTTCCGAGCCACA 5169  
150 GlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGluIleTyr 169  
5170 CAATTGAAAAACCTCGGAACCCCGAAGTTCAATGATTAATCAAGGTCTCGAGTTATA 5229  
170 LysGlySerAspGlyLys-----IleAspAlaPheValSerGlyIleGlyThrGlyThr 188  
5230 CAACATTTGAGAGGAAAGAACTATGATGCGCTTCTGACTGCTGTGTGCTGCTGATGAG 5289  
189 IleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAlaGlyLysLeuIleGlyVal 208  
5290 TTAATCTGCTAGGATTAAGTTCTGAAAAAAGATATCTTAACATCGAATTTGCTGATA 5349  
209 GlnProValGluSerProValLeuSerGlyGlyAspProGlyProHisLysIleGlnGly 228  
5350 GAGCTTAAAGCTTCTCGATTAATGAGCGGTGATGAGCGGTGATGAGCGGTGATTAATCAAGGT 5409  
229 IleGlyAlaGlyPheIleProGlyValLeuGlnValAsnLeuLeuAspGluValGln 248  
5410 TTAGGTCTGATTAATTTTCCAGGACCTTGAATACAGAAATCATGACATATATTAATAA 5469  
249 IleSerSerAspGluAlaIleGlyThrAlaLysLeuLeuAlaLeuGlyGlyLeuPhe 268  
5470 GTAGAAATATATACAGGATGAGAAATGCTGCTGATGAGTTCTTAAGGAGGATTTTA 5529  
269 ValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysAspProGly 288  
5530 GCAGGTATTTCTATCAGGTGCTGCAATTAATGCTCCATTCGAAAAAGGAAAAAGAA--TTA 5586  
289 AsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluGlyIleLeuSer 308  
5587 GGAAGAGGATTAAGACAGTAAACGATGATGCGAGTATGATGATGAGCGCTATATCAACA 5646  
309 ValLeuPhe 311  
5647 CCTTATAT 5655

RESULT 13  
US-09-107-532A-271  
Sequence 271, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A. Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107.532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Denek  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-8277  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 271:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 927 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURES:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...927  
SEQUENCE DESCRIPTION: SEQ ID NO: 271:  
US-09-107-532A-271

Alignment Scores:  
Pred. No.: 1,43e-86 Length: 927  
Score: 813.00 Matches: 171  
Percent Similarity: 70.83% Conservative: 50  
Best Local Similarity: 54.81% Mismatches: 85  
Query Match: 50.09% Indels: 6  
DB: 4 Gaps: 5

US-09-931-457A-31 (1-325) x US-09-107-532A-271 (1-927)

QY 5 ArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeu 24  
DB 1 AGAGAGCGTTTCTATTAACAGTCAACCGAGTTCGATGGGAAAAACGCAATCGCAATTA 60  
QY 25 AsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluPro 44  
DB 61 AATAAATCTGACCGAAGATTCAGCAGATGTTTGTAAAGCTTGAGTTTATATCA 120  
QY 45 CysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGluLysGly 64  
DB 121 GAGGAGAGGTAAAGATCGATTCGTTTAAGTATGATGAAAAAGCAACACATGCT 180  
QY 65 LeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGly 84  
DB 181 CTACTGAACCGAGGTATCTATC---ATTGAACCTACTCTCGAAATCAAGAAATCGCA 237  
QY 85 LeuAlaPheMetAlaAlaIleArgGlyTyrLysLeuIleIleThrMetProAlaSerMet 104  
DB 238 TTGTCATGATGAGGCTTCAAGAGGATTAAGATTAAGTATGATGCTGATGCTGATG 297  
QY 105 SerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuValLeuThrAspPro 124  
DB 298 AGTATCGAAAGAGCGTTTATGATGAAGGTCAGGCGCAGAGTATTTGACTCTGCT 357  
QY 125 AlaLysGlyMetLysGlyAlaValGlnLysAlaGluGluIleLeuAlaLysThrProAsn 144  
DB 358 GCGAGTGTATTTCTGATCAATCAAGAGCAAGACGT---TTGCAAAAGAAAAACGCG 414  
QY 145 AlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThr 164  
DB 415 TACTTTTACATTAATTTGAAAAATGAAGCAATTCCTTGATGATGAGAGCACT 474  
QY 165 GlyProGluIleTyrPhe-----GlySerAspGlyLysIleAspAlaPheValSerGly 182  
DB 475 GGTCTGAGATACATCAACGATTTGAGTAATATGCG---TTGAGCGCGTTTGTGCAAGGT 531



QY 183 IleglythrglythrllethrglyalaglystfyleuylsgluqlasnpProasn 202  
 Db 532 ATCGGTACAGAGAGACGATCATCGCGCTGCGAGAAATTAACGTGTATCCAAA 591  
 QY 203 IleylsleuilegylvalgluProvalgluSerProvalleuSerlyglyLysProglly 222  
 Db 592 ATCGAATTAATCGGGTGAACACAGACAAATCTGCTATTATTCAGAAAGGAAAGCGGA 651  
 QY 223 ProhislyrilegnglylleglyalaglypheilleProgllyvalleuqluvalasnu 242  
 Db 652 CCTCAAAATCCAGAGAAATCGTACAGTTTGTCTTAAACACAGTGAATCTTGTG 711  
 QY 243 LeuapgluvalalaglnlleserSerapglualllegluthrAlalyseuileuAla 262  
 Db 712 TATATTAAGTCTTCTTATTAAGTGAACAGAGAGAAAGCTTACAGAAATGCGG 771  
 QY 263 LeuylsgluqluPhevalgllylleserSerlyalalalalalalalalalalal 282  
 Db 772 CGAAAGAAAGAGATTTAGTAGCATCTCATCTGCGCTCAATGCTGCGCTAA 831  
 QY 283 IleylalyaargProgluasnAlaglylyseuileuileuAlalyPheProSerPhegly 302  
 Db 832 GTCCGAAAGAA--TTAGAAAGAGCAAGAGTATTAGCTGTGTTCCGCAATGCT 888  
 QY 303 GluArglyrleuSerSerValleuPhegluSerVal 314  
 Db 889 GACCTTATTTATCTACAGCCCTTACCAAGAGATA 924

# RESULT 14

US-09-134-001C-493  
 ; Sequence 493, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lym Docette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 493  
 ; LENGTH: 960  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-493

## Alignment Scores:

Pred. No.: 3,99e-85 Length: 960  
 Score: 801.00 Matches: 159  
 Percent Similarity: 72.52% Conservative: 60  
 Best Local Similarity: 52.65% Mismatches: 77  
 Query Match: 49.35% Indels: 6  
 Gaps: 5

US-09-931-457a-31 (1-325) x US-09-134-001C-493 (1-960)

QY 12 ValThrgluLeuilegylserThrProleuVallyrleuasnlyleuAlaapglyCys 31  
 Db 52 GTTACGCAATTAATTTGGGATACACCTGTAGTCAATTAAGAAAGCTTGTATGAT 111  
 QY 32 ValAlaargValAlaAlalyseuileuSerGluProCysSerSerVallylsaparg 51  
 Db 112 GAGCGTATATTATGTTAGTAAATCAAAATCCAGGTGTGCTGCTAAAGATCGT 171  
 QY 52 IleglyrlyrserMerilleAlasplagluqluvalglyleuilethrglyLysSer 71  
 Db 172 ATCGCTTAAGCAATTAAGAAAGCTGAGCGTGAAGGAAATTAACCTGTATACA 231  
 QY 72 ValleuilegluProthrSerlyasnThrglylleglyleuAlapheveAlaAlaAla 91

Db 232 ATC---GTTAGCTTACAGAGTGTACATCGATATAGTCTACATTTGTATATGCTGCC 288  
 QY 92 ArgglyrlyrlyseuilelethrmelProAlaserSerleuqluArgglyle 111  
 Db 289 AAGGGTACAAAGCAAGTTTATACATGCCCTGAACATGAGCCAGAGCGCCGACTTA 348  
 QY 112 LeuLeuAlaPhegllyAlagluLeuValleuThrAspProAlalyseuileuAla 131  
 Db 349 TTAAGACTATATGCTGTAAGTATGATTAATTAACCCAGATCTAGACTATGAAGTGA 408  
 QY 132 ValGlnlysaIagluqluileuAlalyserThrProsnAlalyrile---LeuGln 150  
 Db 409 ATTAAGAAAGCTTAAGATTA-----AAGAAAGACACCGCTATTGAAACCAACAA 462  
 QY 151 PhegluasnProAlaenProlysValhietryoluThrglyrProgllyletrlys 170  
 Db 463 TTCGAAATCCAGCAAAATCCGAATTCATGAACCTTCAATGACCAAGATTAGTTGA 522  
 QY 171 GlySerapglylys---Ileasplaphevalserglylleglythrglythrlle 189  
 Db 523 CAATTGAAGTCAACAAATTAATGATTTTATGCTGTGTAGAACTGGGTGACTTA 582  
 QY 190 ThrglyalaglyrlyrlyrleuylsgluqlasnpProasnleuileuileuileu 209  
 Db 583 TCTGTGTGTGTAAAGTATGAAGAAAGAAATATCCAAATGTGAAATGAGCTATTGAA 642  
 QY 210 ProvalgluSerProvalleuSerglylyrProgllyProhislyrilegngly 229  
 Db 643 CTGAAGCTTCTCAAGTATTAAGCGGTGGAACAGAGCCCTCAATTAATTAAGAGTTG 702  
 QY 230 GlylalyPheilleProgllyvalleuqluvalasnuleuAlalyPhegllyle 249  
 Db 703 GAGACAGTTTCGACTGATCTTAAATTAACAAAGTTTATGACACATCATCAAGATA 762  
 QY 250 SererapglualllegluthrAlalyseuileuileuylsgluqluPheval 269  
 Db 763 GGTATATACTGCTATGATGATGACAGCTGTGCTTGAAGAAAGATATATGACA 822  
 QY 270 GlylleserSerglyalalalalalalalalalalalalyaargProgluasn 289  
 Db 823 GGTATTTATCTGTGTGCGCAATATATGCTCTTCAAAAGCAAAAGAA--TTAGT 879  
 QY 290 AlaelylyseuileuileuAlalyPheProSerPheglyGluArglyrleuSerSerVal 309  
 Db 880 AAGGTAAACAGTTTAAACAGTATTAACCAAGTATGAGGAAAGCTTATTAACACCA 939  
 QY 310 Leuphe 311  
 Db 940 TTATAT 945

## RESULT 15

US-08-961-527-15  
 ; Sequence 15, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527

Mon May 24 08:18:51 2004

us-09-931-457a-31.p2n.rn1

Page 12

FILING DATE: 424  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10993 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-15

Alignment Scores:  
Pred. No.: 1,366-80 Length: 10993  
Score: 777.00 Matches: 162  
Percent Similarity: 70.68% Conservative: 55  
Best Local Similarity: 52.77% Mismatches: 84  
Query Match: 47.87% Indels: 6  
DB: 4 Gaps: 5

US-09-931-457a-31 (1-325) x US-08-961-527-15 (1-10993)

QY 8 ILEALALYASPVATITRNGIULENIIIEGLYVSETRPROLEUVALTYRLEUASNUYLEU 27  
DB 7400 ATTATATACACACTACTGATTTATGCTGTAACACCGATTTGTAACCTTAAACATC 7459  
QY 28 ALASPGIYCYSAVALAARGVALAALALYSLEUIGLUUWETGIUPROCYSSER 47  
DB 7460 GTCCAGAGAGTGTGAGAGCTTATATTAAGCTTAAAGCATTTATCTGATCATCT 7519  
QY 48 VALIYASAPRIGILEGIYTYRSEMERILEALASPLAGIUGIULYSGIYLEUIETHR 67  
DB 7520 GFAAAGACCGTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 7579  
QY 68 PROGLYVSERVALLEUIGLUUPROTHSERGIYASANTHRIGIYILEUALAPHE 87  
DB 7580 CTTGCT---TCTACTATTTGTAAGCAACAGTGAACACCGTATTTGCACTTCAAG 7636  
QY 88 METALALIALAARGIYTYRYSLEUILEIETHECTPROALASERMETSERLEUGLU 107  
DB 7637 GFAAGTCTCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAG 7696  
QY 108 ARGARGILEIIELEUENALAPHEGLYVALAGLUUENVALIENUTHRAPPROALATYSGLY 127  
DB 7697 CGAGTAAATTTATTCAGACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7756  
QY 128 METYSGLIYALAVAGIULYVALAGLUIGLUILEUUALYSTHRPROASNAIATYRTLE 147  
DB 7757 ATGAAGTCTCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAG 7813  
QY 148 LEUENGLINHEGILUASNPPOALASNPPOALASNPPOALASNPPOALASNPPOALASNP 167  
DB 7814 CTTCTTAATTGCAATTCAGACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7873  
QY 168 ILETPLYS-----GIYSEASPGIYLYSILEAPALAPHEVALISERGIYILEGIYTHR 185  
DB 7874 ATACTACTGCTTCTGCTTAAGATGGA---TTAGATGCTTCTGCTGCTGCTGCTGCTGCT 7930  
QY 186 GLYGLYTHRIETHRIGIYALAGIYLYSTYRLEUIGLUIGLUINSENPROASNIIELYSLEU 205  
DB 7931 GGTGGAACGATTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7990  
QY 206 IIEGLYVALGIUPROVALIUSERPROVALIUSERGIYGLIYSPROGIYPROHISLYS 225  
DB 7991 TTTCAGATGAGACAGATGATCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8050

QY 226 IIEGLINGIYILEGIYALAGIYPHEIIEPROGLIYVALLEUGLUVALASNULEUASPGIU 245  
DB 8051 ATTCAAGATATCTGCTGCTGCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8110  
QY 246 VALIYALGINIIESESERASPGLUAIILEGIUTHRALALYSLEUENALALEUIGLU 265  
DB 8111 ATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8170  
QY 266 GLYLEUPHEVALIGIYIIESESERGIYALALALALALALALALALALALALALALAL 285  
DB 8171 GGTCTCTTGTGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8230  
QY 286 ARGPROGLUASNAIIGIYLYSLEUILEVALIYALPHEPROSERPHEGIYGLIYARGTYR 305  
DB 8231 AAA---TTAGTACAGGTAAAGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8287  
QY 306 LEUSERSEVALLEUPHEGLU 312  
DB 8288 CTTCTACAGCACTTATGAA 8308

Search completed: May 22, 2004, 01:43:08  
Job time : 3996 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 21, 2004, 23:43:26 ; Search time 577 Seconds  
(without alignments)  
2559.855 Million cell updates/sec

Title: US-09-931-457A-31

Percent score: 1623  
Sequence: 1 MAVERSGIAXQVTELLIKTP.....ISSVHFSVSRREASQMTREP 325

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODE=frame\_plus.model -US09931457/runac.18052004.121731.20391/app.query.fasta.1.519  
-Q=/cgn2\_1/USPTO\_spool/US09931457 -SUFFIX=p2n.tmpd -MINMATCH=0.1  
-DB=Published Applications NA -QMT=faetap -SUFFIX=p2n.tmpd -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USR=US09931457 @CCN 1.1.57 @runac.18052004.121731.20391  
-NCPD=6 -ICPD=3 -NO WAMP -LANG=JQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100 -XGAPEXT=0.5  
-PONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query Match Length DB ID Description

1	1623	100.0	1362	9	US-09-931-457A-30	Sequence 30, Appl
2	1622	99.9	1407	13	US-10-424-599-130460	Sequence 130460, A
3	1488	91.7	3075	13	US-10-424-599-55164	Sequence 55164, A
4	1389	85.6	1319	13	US-10-425-114-25435	Sequence 25435, A
5	1381	85.1	1308	13	US-10-425-114-19467	Sequence 19467, A
6	1381	85.1	1403	13	US-10-425-114-1445	Sequence 1445, Ap
7	1380	85.0	1312	13	US-10-425-114-23288	Sequence 23288, A
8	1375	84.7	1301	13	US-10-425-114-23608	Sequence 23608, A
9	1374	84.7	966	9	US-09-938-842A-893	Sequence 893, App
10	1374	84.7	966	11	US-09-938-842A-893	Sequence 893, App
11	1317	81.1	1522	9	US-09-931-457A-63	Sequence 63, Appl
12	1316	81.1	1404	13	US-10-424-599-55160	Sequence 55160, A
13	1309	80.7	1279	13	US-10-425-114-22525	Sequence 22525, A
14	1279	78.8	1490	13	US-10-425-114-34372	Sequence 34372, A
15	1279	78.8	1491	13	US-10-425-114-34527	Sequence 34527, A
16	1275	78.6	1508	9	US-09-931-457A-61	Sequence 61, Appl
17	1275	78.6	1509	9	US-10-425-114-30493	Sequence 30493, A
18	1255	77.3	1329	9	US-09-938-842A-99	Sequence 99, Appl
19	1255	77.3	1329	9	US-09-938-842A-99	Sequence 99, Appl
20	1234	76.0	1603	13	US-10-424-599-61234	Sequence 61234, A
21	1192	75.4	2307	13	US-10-425-114-7548	Sequence 7548, Ap
22	1192	73.4	972	9	US-09-938-842A-495	Sequence 495, App
23	1192	73.4	1472	11	US-09-938-842A-495	Sequence 495, App
24	1179	72.6	975	10	US-09-905-290A-3	Sequence 3, Appl
25	1179	72.6	975	10	US-10-424-599-53377	Sequence 53377, A
26	1079	66.5	1563	13	US-10-282-122A-12335	Sequence 12335, A
27	968.5	59.7	945	13	US-10-282-122A-26520	Sequence 26520, A
28	949.5	58.5	930	13	US-10-282-122A-26508	Sequence 26508, A
29	949.5	58.5	933	13	US-10-282-122A-26508	Sequence 26508, A
30	947.5	58.4	1925	13	US-10-424-599-18037	Sequence 18037, A
31	941.5	58.0	930	13	US-10-282-122A-26080	Sequence 26080, A
32	941.5	58.0	1086	13	US-10-282-122A-27578	Sequence 27578, A
33	940.5	57.9	933	13	US-10-282-122A-26080	Sequence 26080, A
34	917	56.5	927	13	US-10-282-122A-26080	Sequence 26080, A
35	917	56.5	100848	13	US-10-672-787-39	Sequence 39, Appl
36	916.5	56.5	927	13	US-10-282-122A-16019	Sequence 16019, A
37	912	56.2	933	13	US-10-282-122A-29654	Sequence 29654, A
38	897	55.3	588	16	US-10-341-961A-271	Sequence 271, App
39	896.5	55.2	930	13	US-10-282-122A-15648	Sequence 15648, A
40	882	54.3	969	13	US-10-282-122A-14010	Sequence 14010, A
41	879	54.2	951	13	US-10-282-122A-32551	Sequence 32551, A
42	870	53.6	996	13	US-10-282-122A-8820	Sequence 8820, Ap
43	865.5	53.3	972	13	US-10-282-122A-38798	Sequence 38798, A
44	865.5	53.3	972	13	US-10-282-122A-39402	Sequence 39402, A
45	863.5	53.2	969	13	US-10-282-122A-19471	Sequence 19471, A

# ALIGNMENTS

RESULT 1  
US-09-931-457A-30  
Sequence 30, Application US/09931457A  
Patent No. US20020157132A1  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
FILE REFERENCE: BB1116 US CIP  
CURRENT FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: US/09/931, 457A  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065, 385  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: 60/049, 406  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 30  
LENGTH: 1362  
TYPE: DNA  
ORGANISM: Glycine max

US-09-931-457a-30

Alignment Scores:

Pred. No.:	1,84e-178	Length:	1362
Score:	1623.00	Matches:	325
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-931-457a-31 (1-325) x US-09-931-457a-30 (1-1362)

```

QY 1 MetAlValGluArgSerGlyIleAlaIleAspValThrGluLeuIleGlyLysThrPro 20
DB 90 ATGGCTGTTGAAGAGCTTCCGAAATTCCTCAAGATTCGATTTGGTAAACCCCA 149
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 150 TTAGTATATCTAATAATAAAGCTTCGAGATGTTGTGTGCCCGGCTTCTCTAAACCGAG 209
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
DB 210 TTATGAGAGCCAGCTCTAGTGTGAAGACAGATTCGGTATAGTATGCTGATTCGA 269
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValIleuIleGluProThrSerGlyAsn 80
DB 270 GAAGAGAGGAGCTTATCAACCTGGAAAGAGTGTCTCATAGCCCAACAGTGTAAAT 329
QY 81 ThrGlyTyrIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
DB 330 ACTGGCATTTGATTAAGCTTCATGAGCGACGCGGATTAACAGCTCAATTAACATG 389
QY 101 ProAlaSerSerSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
DB 390 CCGCTCTTATAGAGCTCTTGAGAAAGATCATTTCTATAGCTTTTGAGCTGAGTGGT 449
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValAlaGluLysAlaGluIleLeuAla 140
DB 450 CTGACAGATCTCTGCTAAGGAAATGAAGGTGCTTCAAGAGGCTCAAGAGATTTGGCT 509
QY 141 LysThrProAsnAlaTyrIleLeuGlnAlaPheGluAsnProAlaAsnProLysValHis 160
DB 510 AAGACGCCCAATCTCAATCTTCAACATTTGAAACCTCCCAATCCCAAGCTTCAAT 569
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180
DB 570 TATGAAACCACTGCTCCAGATATGAAAGGCTCCGATGCGAAATTTGATGCAATTTGTT 629
QY 181 SerGlyTyrIleGlyThrGlyTyrThrIleThrGlyAlaGlyLysTyrLeuLysGluIleAsn 200
DB 630 TCGGGATAGGCACTGGGTGTAACATTAACAGTCTGTAAGAAATTTCTTAAGAGCGAAT 689
QY 201 ProAsnIleLysLeuIleGlyValAlaGluProValGluSerProValLeuSerGlyLys 220
DB 690 CCGAATATAAAGCTGATGTTGTTGGAACAGTTGAAAGTCACTGCTCTCAAGAGAAAG 749
QY 221 ThrGlyProHisTyrIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240
DB 750 CCGGTTCACACAGAGATTCAGGAGATTTGCTGTTTATCCCTGATGCTTGAAGATC 809
QY 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
DB 810 AACTCTTCTGATGAAGTTGTTCAAAATATCAAGATGATCAACATTAAGAACTGCAAGCTT 869
QY 261 LeuAlaLeuLysGluIleLeuPheValAlaGlyIleSerSerGlyValAlaAlaAlaAlaAla 280
DB 870 CTTCGGCTTAAAGAGGCTTATTTGGTGAATATCTTCGAGAGCTGAGTGTGCTGCT 929
QY 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300
DB 930 TTTCAGATTCGAAGAAAGCAAGAAATGCCGGAAGCTTATTTGCGCGTTTTCACAG 989
QY 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320

```

```

DB 990 TTCCGGAGAGAGTACTCTCTCTCTGCTATTGAGTCAAGTGAAGACGGAAGCTAAAGC 1049
QY 321 MetThrPheGluPro 325
DB 1050 ATGACTTTTGAAGCCC 1064

RESULT 2
US-10-424-599-130460
/ Sequence 130460, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 130460
/ LENGTH: 1407
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_88813C.1
US-10-424-599-130460

Alignment Scores:
Pred. No.: 2.52e-178 Length: 1407
Score: 1622.00 Matches: 324
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.69% Mismatches: 0
Query Match: 99.94% Indels: 0
Gaps: 0

US-09-931-457a-31 (1-325) x US-10-424-599-130460 (1-1407)
QY 1 MetAlaValGluArgSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
DB 126 ATGGCTGTTGAAGAGCTTCCGAAATTCCTCAAGATTCGATTTGGTAAACCCCA 185
QY 21 LeuValTyrLeuAsnLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
DB 186 TTAGTATATCTAATAATAAAGCTTCGAGATGTTGTGTGCCCGGCTTCTCTAAACCGCA 245
QY 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
DB 246 TTGATGAGAGCCATGCTCTAGTGTGAAGACAGATTCGATTAATGATTCGATTCGA 305
QY 61 GluGluLysGlyLeuIleThrProGlyLysSerValIleuIleGluProThrSerGlyAsn 80
DB 306 GAAGAGAGGAGCTTATCAACCTGGAAGAGTGTCTCATATGACCCCAAGAGTGAAT 365
QY 306 GAAGAGAGGAGCTTATCAACCTGGAAGAGTGTCTCATATGACCCCAAGAGTGAAT 365
QY 81 ThrGlyTyrIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
DB 366 ACTGCAATTCGATTAAGCTTCAATGCAACAGCGGGTTCAACACTCATTAATCAATG 425
QY 101 ProAlaSerSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValAlaGluVal 120
DB 426 CCGCTCTTATGAGCTTGAAGAGAAATCATTTCAATAGCTTTTGAGAGCTGAGTGTGTT 485
QY 121 LeuThrAspProAlaLysGlyMetLysGlyAlaValAlaGlnLysAlaGluIleLeuAla 140
DB 486 CTGACAGATCTCTCAAGGAGATGAAGTCTGTTCAAGAGCTGGAAGATTTGGCT 545
QY 141 LysThrProAsnAlaTyrIleLeuGlnAlaPheGluAsnProAlaAsnProLysValHis 160
DB 546 AAGAGGCCCAATGCTTACATATCTTCAACATTTGAAACCTTCGCAATCCCAAGTTCAT 605
QY 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180

```



Mon May 24 08:18:52 2004

us-09-931-457a-31.p2n.rnpb

**Page 4**

```

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25435
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3900-005-C7_FLI
US-10-425-114-25435

```

DB 401 AGAGGATCATATGAGGCTTTGGTGTGACCTTCTTACTGACCCACTCTGGGA 460  
 QY 128 MetLysGlyAlaValAGlnLysAlaGlnLysLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 461 ATGAAGAGCTGTCAAGAAAGCCGAGAGATACAGCAAGAACACCCAACTCGTATC 520  
 QY 148 LeuGlnGlnPheGlnPheProAlaAsnProLysValHisTyrGlnLysThrGlyProGln 167  
 DB 521 CTCAACATTGAAATCCAGCTAACCCAAAGATTCATAGAGACTACGCGCCCTGAA 580  
 QY 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 187  
 DB 581 ATCTGAAAGCTACAGCAGGAAATTCCTGGCTGTATCTGTATCTGGATTCGGAGAGGT 640  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGlnLysProAsnIleLysLeuIleGly 207  
 DB 641 ACCATCAACCGGTACTGGCGCATCTCAGAGAGCAAAATCTTAATGTCAAGCTCTATGT 700  
 QY 208 ValGlnProValGlnSerProValLeuSerGlyLysProGlyProHisLysIleGln 227  
 DB 701 GTGAGCCAGTGAAGAGTCTGTTTGAATGTGAAATCCTGACACACACAAATTC 760  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGlnValAsnLeuLeuAspGlyVal 247  
 DB 761 GGAATGTGAGCTGTTTATCCCTGAGAGCTTGATGTGATCTTGTATGATGAGACTCT 820  
 QY 248 GlnIleSerSerAspGlnAlaIleGlnThrAlaLysLeuAlaLysGlnLysLeu 267  
 DB 821 CAGTTTCAAGTGAATGAAGCTATTTAGAGCTCCAGAGCTCTGTCTGTAAGAGAGGTG 880  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 DB 881 TTGTTTGAATCTTCTTGTGAGAGCTGACCTGCGGAGTTAGCTTGTCTTAAGAGCCA 940  
 QY 288 GlnAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlnTyrLeuSer 307  
 DB 941 GAAAGCCCGGAAAGCTATTTGTGTCTGTCTCCAGAGCTTCGCGAGCGCTTACTCTCA 1000  
 QY 308 SerValLeuPheGlnSerValArgArgGlnAlaGlnSerMetThrPheGlnPro 325  
 DB 1001 TCGGTCTGTTCACGTCCATCAAGAGAGCGGAAAGCAATGCTGTGAGGCC 1054

RESULT 6

US-10-425-114-1445  
 : Sequence 1445, Application US/10425114  
 : Publication No. US2004003488A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Liu, Jindong  
 : APPLICANT: Zhou, Yihua  
 : APPLICANT: Kovalic, David K.  
 : APPLICANT: Screen, Steven E  
 : APPLICANT: Tabaska, Jack E  
 : APPLICANT: Cao, Yongwei  
 : TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 : TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 : FILE REFERENCE: 38-21(5313)B  
 : CURRENT APPLICATION NUMBER: US/10/425,114  
 : NUMBER OF SEQ ID NOS: 2003-04-28  
 : SEQ ID NO 1445  
 : LENGTH: 1403  
 : TYPE: DNA  
 : ORGANISM: Zea mays  
 : FEATURES:  
 : OTHER INFORMATION: Clone ID: 700153393\_FLI  
 : US-10-425-114-1445

Alignment Scores:  
 Pred. No.: 2.57e-150 Length: 1403  
 Score: 1381.00 Matches: 269  
 Percent Similarity: 91.51% Conservative: 22  
 Best Local Similarity: 84.59% Mismatches: 27  
 Query Match: 85.09% Indels: 0

DB: 13 Gaps: 0

US-09-931-457a-31 (1-325) x US-10-425-114-1445 (1-1403)

QY 8 IleAlaLysAspValThrGlnLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27  
 DB 196 ATGCCAGAGACTCTACCGAGTTGATCGGAAAGACCGCGTGTATCTCAACCAAGGTG 255  
 QY 28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGlnLeuMetGlnProCysSerSer 47  
 DB 256 ACCGATGGGTGCTCCGCGCCGCTGCGCCCAAGCTGAGTCCATGAGGCCCTGCTCAGC 315  
 QY 48 ValLysAspArgIleGlyTyrSerMetIleAlaAspAlaGlnLysGlyLysLeuThr 67  
 DB 316 GTCAAGATGAGATTGGCTACAGCATATACAGGACGAGAGAGAAAGGCTGATTACT 375  
 QY 68 ProGlyLysSerValLeuIleGlnProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87  
 DB 376 CCAAGCGTGAAGTGTGATTGATGAACCAACTAGCGGCAACAGGCAATTGACTGGCTTT 435  
 QY 88 MetAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGln 107  
 DB 436 ATGCTGTCTCCAGAGGCTACAACTTACCTCACTCAATGCTCTCTCATGACATGAG 495  
 QY 108 ArgArgIleIleLeuLeuAlaPheGlyAlaGlnLeuValLeuThrAspProAlaLysGly 127  
 DB 496 AGAGGATCATATTTGAAGCTTTTGGTGTGAACTTGTCTTACTGACCACTCTGGGA 555  
 QY 128 MetLysGlyAlaValAGlnLysAlaGlnLysLeuAlaLysThrProAsnAlaTyrIle 147  
 DB 556 ATGAAGAGACTCTTCAAGAAAGCGGAGATACAGCAAGAACACCCAACTGATATC 615  
 QY 148 LeuGlnGlnPheGlnPheProAlaAsnProLysValHisTyrGlnLysThrGlyProGln 167  
 DB 616 CTTCACAAATTTGAAATCCAGCTTAACCCAAATTCATCTATGACTATGAGGCGTGA 675  
 QY 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGly 187  
 DB 676 ATCTGAAAGCTACAGCAGGAAATTCGCGCTTGTATCTGTATCTGAGTCCGGAAGAGGT 735  
 QY 188 ThrIleThrGlyAlaGlyLysTyrLeuLysGlnLysProAsnIleLysLeuIleGly 207  
 DB 736 ACCATCAACCGGTACTGGCGCATCTCAGAGAGCAAAATCTTAATGCAAGCTCTATGT 795  
 QY 208 ValGlnProValGlnSerProValLeuSerGlyLysProGlyProHisLysIleGln 227  
 DB 796 GTGAGCCAGTGAAGAGTGTGTTTGAATGTGGAAGAACTGAGACACCAAGATTCAA 855  
 QY 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGlnValAsnLeuLeuAspGlyVal 247  
 DB 856 GGAATGTGAGCTGTTTATCCCTGAGAGCTTGAATGTATCTTGAATGAGACTCTA 915  
 QY 248 GlnIleSerSerAspGlnAlaIleGlnThrAlaLysLeuAlaLysGlnLysLeu 267  
 DB 916 CAGTTTCAAGTGAATGAAGCTATTTAGAGCTGCGCAAGCTCTTCTCTGAAAGAGGTG 975  
 QY 268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287  
 DB 976 TTGTTTGAATCTTCTTGTGAGAGCTGAGCTGCGGCAAGTTAGCTTGAAGAGCCA 1035  
 QY 288 GlnAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGlnTyrLeuSer 307  
 DB 1036 GAAAGCCCGGAAAGCTATTTGTGTCTGTCTCCAGAGCTTCGCGAGCGCTACTCTCA 1095  
 QY 308 SerValLeuPheGlnSerValArgArgGlnAlaGlnSerMetThrPheGlnPro 325  
 DB 1096 TCGGTCTGTTCACGTCCATCAAGAGAGAGCGGAAAGCAATGCTGTGAGGCC 1149

RESULT 7

US-10-425-114-23298  
 : Sequence 23298, Application US/10425114  
 : Publication No. US2004003488A1  
 : GENERAL INFORMATION:

```

# APPLICANT: Liu, Jindong
# APPLICANT: Zhou, Yihua
# APPLICANT: Kovalic, David K.
# APPLICANT: Screen, Steven E.
# APPLICANT: Tabaska, Jack B
# APPLICANT: Cao, Yongwei
# TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
# TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
# FILE REFERENCE: 38-21(53313)B
# CURRENT APPLICATION NUMBER: US/10/425,114
# CURRENT FILING DATE: 2003-04-28
# NUMBER OF SEQ ID NOS: 73128
# SEQ ID NO 23298
# LENGTH: 1312
# TYPE: DNA
# ORGANISM: Zea mays
# FEATURE:
# OTHER INFORMATION: Clone ID: LIB3597-017-F4_P1
US-10-425-114-23298

```



```

Db 334 ATGGCTGCTCCAGAGGCTTACCAACTTACCTACCAATGCTGCTCATATAGCATGAG 393
Qy 108 ArgArgIleIleLeuLeuAlaPheGlyValAGluLeuValLeuThrAspProAlaLysGly 127
Db 394 AGGAGATCATATATGAGAGGCTTTTGTGCTGATCTGCTTACCTGACCCACTCTGGGA 453
Qy 128 MetIysGlyAlaValAGluLeuValAGluGluIleLeuAlaLysThrProAlaLysIle 147
Db 454 ATGAAAGAGCTGTCAAGAAAGCGAAGAAATACAGCAAGAAAGCAACCACTCTGACATT 513
Qy 148 LeuGlnGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167
Db 514 CTTCACAAATTTGAAATCCAGCTTACCCAAAGATTCCTGTAGACTACTAGGGCTGAA 573
Qy 168 IleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187
Db 574 ATCTGAAAGCTACAGCAGAAATATGATGCTTGTATCTGTGTGCGGAGCAGAGGT 633
Qy 188 ThrIleThrGlyValAGlyLysTyrLeuLysGluGlnAsnProAlaLysLeuIleGly 207
Db 634 ACCATACCCGATCTGCGGCGATCTCTGACAGAGCAAAATCTTATGTCAGACTCTATGCT 693
Qy 208 ValGluProValAGluSerProValLeuSerGlyLysProGlyProGlyLysIleGln 227
Db 694 GTGGAGCCAGTGGAGAGTGTGTTTGAATGTGGAAACCTGGACCAAGATTCAA 753
Qy 228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValAsnLeuLeuAspGluVal 247
Db 754 GGAATGAGCTGCTTATCTCTGAGCTTGTGATGTTGATCTCTTGAAGAACTCTA 813
Qy 248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267
Db 814 CAGGTTCAAGTGAAGTGAAGTATGACAGCTGCAAGGCTCTTGGCTTGAAGAAAGGCTG 873
Qy 268 PheValGlyIleSerSerGlyValAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287
Db 874 TTGGTTGAATCTCTTCTGTGCTGACCTGACCTGCGGAGTGAAGTCTTGAAGGCGCA 933
Qy 288 GluAsnAlaGlyLysLeuIleValAlaValPheProSerPheGlyGluArgTyrLeuSer 307
Db 934 GAAAGCGCCGGAAGCTATTTGTTGTCTGCTTCCGACCTTCCGGAAGCTCACTCTCA 993
Qy 308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325
Db 994 TCGGTGTGTTCCATCATCATCAAGAAAGAGGAGACATGCTGTGAGGCC 1047

```

RESULT 9  
 US-09-938-842A-893  
 Sequence 893, Application US/09938842A  
 Patent No. US20020160378A1  
 GENERAL INFORMATION:  
 APPLICANT: Harper, Jeff  
 APPLICANT: Kreps, Joel  
 APPLICANT: Wang, Xun  
 APPLICANT: Zhu, Tong  
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 TITLE OF INVENTION: SAME, AND METHODS OF USE  
 FILE REFERENCE: S01P1300-3  
 CURRENT APPLICATION NUMBER: US/09/938,842A  
 CURRENT FILING DATE: 2001-08-24  
 PRIOR APPLICATION NUMBER: US 60/227,866  
 PRIOR FILING DATE: 2000-08-24  
 PRIOR APPLICATION NUMBER: US 60/264,647  
 PRIOR FILING DATE: 2001-01-16  
 PRIOR APPLICATION NUMBER: US 60/300,111  
 PRIOR FILING DATE: 2001-06-22  
 NUMBER OF SEQ ID NOS: 5379  
 SEQ ID NO 893  
 LENGTH: 966  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-893

Alignment Scores:  
 Pred. No.: 9,75e-150 Length: 966  
 Score: 1374.00 Matches: 267  
 Percent Similarity: 91.08% Conservative: 29  
 Best Local Similarity: 82.15% Mismatches: 25  
 Query Match: 84.65% Indels: 4  
 Gaps: 1

US-09-931-457A-31 (1-325) x US-09-938-842A-893 (1-966)

```

Qy 1 MetAlaValAGluSerGlyIleAlaLysAspValThrGluLeuIleGlyLysThrPro 20
Db 1 ATGGCTTCAATGAGCTCTTAAATTCGCTTAAAGATGTAAGTAACTAATAGGAAACATCCA 60
Qy 21 LeuValTyrLeuLeuLysLeuAlaAspGlyCysValAlaArgValAlaAlaLysLeuGlu 40
Db 61 TTATGTATCTTAAACAAAGAGTGAAGATGTTGTGTCAATGTTGCTGCTTACGCTTACG 120
Qy 41 LeuMetGluProCysSerSerValLysAspArgIleGlyTyrSerMetIleAlaAspAla 60
Db 121 ATGATGAAACCAATTTCTTACGCTCAAGACAGGATTTGTTATGATGATGCTGATGCA 180
Qy 61 GluGluLysGlyLeuIleThrProGlyLysSerValLeuIleGluProThrSerGlyAsn 80
Db 181 GAAGCTAAGAGACTTATTAACCGGAGAGAGTGTGATGATGAGCCAAAGAGGAGAC 240
Qy 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMet 100
Db 241 ACTGAGTGGTTTACATTTTACGACGCTCCAAAGGCTTAAAGCTTGTATTAACATG 300
Qy 101 ProLaseMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyAlaGluLeuVal 120
Db 301 CAGGCTCAATGAGTATAGAGAGAGATCAATCTCTTACCTTTGAGCTGAGCTTATTT 360
Qy 121 LeuThrAspProAlaLysGlyMetLysGlyValAlaGluLysAlaGluGluIleLeuAla 140
Db 361 CTATCTATCTCGCTTAAAGTATGAAAGTGTCTGTTCAAAGCGCGAGAGATTTGGCG 420
Qy 141 LysThrProAsnAlaTyrIleLeuGlnGlnPheGluAsnProAlaAsnProLysValHis 160
Db 421 AAAACACTAATGTTATATGCTTCAACAGTTGAGATCTGCTAATCCAAAGATCCAT 480
Qy 161 TyrGluThrThrGlyProGluIleTyrLysGlySerAspGlyLysIleAspAlaPheVal 180
Db 481 TATAGAGCTACTGACCTGAGATATGAAAGTGTCTGTGAAAGTGAAGGCTTTGTT 540
Qy 181 SerGlyIleGlyThrGlyIleThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsn 200
Db 541 TCTGTATGCTTATGCTG-----GGTCTGGGAAGTATCTCAAGAAACAGAAC 588
Qy 201 ProAsnIleLysLeuIleGlyValAGluProValAGluSerProValLeuSerGlyLys 220
Db 589 ACAAAACATAAGCTGTATGTGTGACCTGTGTAAAGCCCTTATCTGTCTGGTGAAG 648
Qy 221 ProGlyProHisLysIleGlnGlyIleGlyValAGlyPheIlePheProGlyValLeuGluVal 240
Db 649 CCAAGTCCCAAAATTCAGAGTATGAGTGTGCTTTTATCCAGGCATTTTGAGTGT 708
Qy 241 AsnLeuLeuAspGluValAlaGlnIleSerSerAspGluAlaIleGluThrAlaLysLeu 260
Db 709 GATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
Qy 261 LeuAlaLeuLysGlyLysPheValAGlyLysSerGlyAlaAlaAlaAlaAlaAlaAla 280
Db 769 CTTCCTGAGAGAGAGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 828
Qy 281 PheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaAlaPheProSer 300
Db 829 ATCAAACTGCAAGAGAGGCAAGAAATGCTGGGAAGCTTATGATGATGATGATGATGAT 888
Qy 301 PheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgArgGluAlaGluSer 320

```

DB 889 TTGGGGGAAAGATATCTATGACGTGATCTGTGATGACGAGCAAGAAAGGACGAGACT 948

QY 321 MetThPheGluPro 325

DB 949 ATGACCTTTGAGCCC 963

RESULT 10

US-09-931-457a-893

Sequence 893, Application US/09938842A

Publication No. US2004009476A9

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPT300-3

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 893

LENGTH: 966

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-931-457a-893

Alignment Scores:

Pred. No.:	9,75e-150	Length:	966
Score:	1374.00	Matches:	267
Percent Similarity:	91.08%	Conservative:	29
Best Local Similarity:	82.15%	Mismatches:	25
Query Match:	84.66%	Indels:	4
DB:	11	Gaps:	1

US-09-931-457a-31 (1-325) x US-09-931-842a-893 (1-966)

QY 1 MetAlValGluArgSerGlyIleAlaIysAspValThGluLeuIleGlyIysThrPro 20

DB 1 ATGGCTTCACTGCTCTTAATTCGCTTAAGATGTTAAGTAACTTAATAGGAACTCA 60

QY 21 LeuValTyrLeuAsnIysLeuAlaAspGlyCysValAlaArgValAlaIysLeuGlu 40

DB 61 TAGTGATATCTAACAAGAGTAAAGATGTTGTCATGTCCTCTCTAAGCTTGAG 120

QY 41 LeuMetGluProCysSerSerValIysAspArgIleGlyTyrSerMetIleAlaAspAla 60

DB 121 ATGATGAACCAATGTTCTAGCGCTCAAGACAGATGGTATAGTATGATGCTGAAGCA 180

QY 61 GluGluIysGlyLeuIleThrProGlyIysSerValIleuIleGluProThrSerGlyAsn 80

DB 181 GAAGCTAAAGGACTTTTAACCGGAGAGAGTGTTGATTAAGCAACACTGGGAAC 240

QY 81 ThrGlyIleGlyLeuAlaPheMetAlaAlaIleArgGlyTyrTyrSerMetIleThrMet 100

DB 241 ACTGAGTGGTGTAGATTACTGACGCTGCAAGAGGCTATAGCTGTATTAACATAG 300

QY 101 ProAlaSerMetSerLeuGluArgIleIleLeuIleAlaPheGlyAlaGluLeuVal 120

DB 301 CCGAGCTTCATATGATATAGAGAGATCACTCTCTGAGCTTTGAGAGCTGATGATT 360

QY 121 LeuThrAspProAlaIysGlyMetIysGlyAlaValAlaGlnIleGluIleLeuAla 140

DB 361 CTACTGATCCGGCTAAAGGTATGAAGGCTGTTGCAAGGCGGAGGAGATTTGGCG 420

QY 141 LysThrProAsnAlaTyrIleLeuGlnIlePheGluAsnProAlaAsnProLysValHis 160

DB 421 AAAACACCTAATGTTATATGCTTCAACAGTTTGAGAAATCCGTAATCCAAAGATCCAT 480

QY 161 TyrGluThrThrGlyProGluIleThrLysGlySerAspGlyIysIleAspAlaPheVal 180

DB 481 TAGGAGACTACTGAGACTGAGATATAGAAAGGTTCTGGTGAAGAGTGAAGGCTTTGTT 540

QY 181 SerGlyIleGlyTyrGlyGlyThrIleThrGlyAlaGlyIysTyrLeuIysGluIleAsn 200

DB 541 TGTGATATTTGACTAGT-----GGTCTGGGAAGTATCTCAAGAAACAGAAC 568

QY 201 ProAsnIleIysLeuIleGlyValGluProValGluSerProValLeuSerGlyIys 220

DB 589 ACAACATATAACCTTATAGTGTCAAGCTGTGTAAGCCCTATTTCTGCGGTGAAG 648

QY 221 ProGlyProHisIysIleGlnGlyIleGlyAlaGlyPheIleProGlyValLeuGluVal 240

DB 649 CCGGTCCTCCCAAAATTCAGAGTATAGTCTGCTGTTTATTCAGGCATTTGGATGTT 708

QY 241 AsnLeuIysAspGluValValGlnIleSerSerAspGluAlaIleGluThrAlaIysLeu 260

DB 709 GATCTTAATAGATGAAGTGTTCAGGTTTCAAGTGAAGATCATGACATGCAAGGCTT 768

QY 261 LeuAlaIysGlyGluIlePheValGlyIleSerSerGlyAlaAlaAlaAlaAla 280

DB 769 CTGGCTTAGGAGAGAGGCTGTTAGTGAATCTATCTGAGCAGCAGCTACTGACACA 828

QY 281 PheGlnIleAlaIysArgProGluAsnIleGlyIysLeuIleValAlaIysPheProSer 300

DB 829 ATCAAACTTGCAAAAGAGGCCAGAAATGCTGGAAAGCTTAATGTGCGGTGTTCCGAGT 888

QY 301 PheGlyIysArgTyrLeuSerSerValLeuPheGluSerValArgArgIleGluIle 320

DB 889 TTGGGGGAAAGATATCTATGACGTGATCTGTGATGACGAGCAAGAAAGGACGAGACT 948

QY 321 MetThPheGluPro 325

DB 949 ATGACCTTTGAGCCC 963

RESULT 11

US-09-931-457a-63

Sequence 63, Application US/09931457A

Patent No. US20020157132A1

GENERAL INFORMATION:

APPLICANT: Falco, S. Carl

APPLICANT: Allen, Stephen M.

TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes

FILE REFERENCE: B1116 US CIP

CURRENT APPLICATION NUMBER: US/09/931,457A

PRIOR FILING DATE: 2002-02-22

PRIOR APPLICATION NUMBER: 09/424,976

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: 60/065,385

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/049,406

NUMBER OF SEQ ID NOS: 72

SOFTWARE: Microsoft Office 97

SEQ ID NO 63

LENGTH: 1522

TYPE: DNA

ORGANISM: Oryza sativa

US-09-931-457a-63

Alignment Scores:

Pred. No.:	7,94e-143	Length:	1522
Score:	1317.00 <td>Matches:</td> <td>250 </td>	Matches:	250
Percent Similarity:	89.47% <td>Conservative:</td> <td>39 </td>	Conservative:	39
Best Local Similarity:	77.40% <td>Mismatches:</td> <td>34 </td>	Mismatches:	34
Query Match:	81.15% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	9 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-09-931-457a-31 (1-325) x US-09-931-457a-63 (1-1522)

3 ValGluArgSerGlyIleAlaIleAspValThrGluLeuIleGlyIleThrProLeuVal 22  
277 GTCCAGGCGCTCAACATCGCCGACGACGCTCACCCAGCTCATCGGCAACATCCATCGTGA 336  
23 TyrLeuAsnIleuAlaAspGlyCysValAlaArgValAlaAlaIleLeuGluLeuMet 42  
337 TATCTCAACAACTGCTCAAGGAGTGTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 396  
43 GluProCysSerSerValIleAspArgIleGlyIleSerMetIleAlaAspAlaGluGlu 62  
397 GAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456  
63 LysGlyIleuIleThrProGlyIleSerValLeuIleGluProThrSerGlyAsnThrGly 82  
457 AAAGGCTGATTAACCTCGAAGAAAGTGTGCTGCAACCAACAAAGTGAATACAGGC 516  
83 IleGlyLeuAlaPheMetAlaAlaAlaArgGlyTyrIleLeuIleIleThrMetProAla 102  
517 ATTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576  
103 SerMetSerLeuGluArgGlyIleIleLeuAlaPheGlyAlaGluLeuValLeuThr 122  
577 TCAATGACCATGAGAGAGAGATTCTACTCAAGCTTGGCGCTGAACTTGTCTTACT 636  
123 AspProAlaIleGlyMetIleGlyValAlaGluIleGlyIleGluIleLeuAlaIleThr 142  
637 GATGCGGCAAGGAGATGAAGGAGCTGTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 696  
143 ProAsnAlaIleIleLeuGluIlePheGluAlaAsnProAlaAsnProIleValIleThrGlu 162  
697 CCGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756  
163 ThrThrGlyProGluIleThrIlePheGlySerAspGlyIleIleAspAlaPheValSerGly 182  
757 ACTACTGGGCGAAGATCTGAGGAGATTCTAAGGAGAGGAGGAGTATTAATGATGCTGGA 816  
183 IleGlyThrGlyIleThrIleThrIleGlyAlaGlyIleIleLeuIleGluIleAsnProAsn 202  
817 ATTGGAACAGGTGCAACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876  
203 IleIleLeuIleGlyValGluProValIleIleSerProValIleSerGlyIleIleProGly 222  
877 ATTAAGGCTTATGATTAAGGCTTCTGAGGATTAACATACCTCTGAGGAGAAACCTGGCC 936  
223 ProIleIleIleGlyIleIleGlyValIleGlyPheIleProGlyValIleGluValIleAsnLeu 242  
937 CCACTAGATTCAGAGCATTTGGGCAAGATTGTTCCAGAGAACTTGATAGTGAAGTT 996  
243 LeuAspGluValIleGluIleSerSerAspGluAlaIleGluThrAlaIleLeuAla 262  
997 CTCGATGAGATTCAGATTAATCTGATGATGAGCTGTTGAGACAGCAACCAATTCGCT 1056  
263 LeuIleGluIleIlePheValIleIleSerSerGlyAlaAlaAlaAlaAlaAlaPheGlu 282  
1057 CTTCAAGAAAGATTAAGTGTGAATTCATCTGGGAGCAAGCAAGCAAGCTGCCATTAA 1116  
283 IleAlaIleArgProGluIleAlaIleGlyIleIleValAlaAlaPheProSerPheGly 302  
1117 GTTCAAAAAAGCAAGAAATGCTGGAAGTGTGAGTGTGCTTCCAAAGCTTTGGT 1176  
303 GluIleGlyIleIleSerSerValIlePheGluSerValaArgGluAlaGluSerMetThr 322  
1177 GAGAGGTACCTTCACTATCTTTTTCATGCTGATTAAGAGAAAGTGTGAAGTGTGCA 1236  
323 PheGluPro 325  
1237 CCTGAACA 1245

APPLICANT: La Rosa Thomas J  
APPLICANT: Kovacic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
NUMBER OF FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 55160  
LENGTH: 1404  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20820C.1  
US-10-424-599-55160  
Alignment Scores:  
Pred. No.: 9,246-143 Length: 1404  
Score: 1316.00 Matches: 273  
Percent Similarity: 89.67% Conservative: 22  
Best Local Similarity: 82.98% Mismatches: 27  
Query Match: 81.08% Indels: 8  
DB: 13 Gaps: 3  
US-09-931-457a-31 (1-325) x US-10-424-599-55160 (1-1404)  
QY 1 MetAlaValGluArgSerGlyIleAlaIleAspValThrGluLeuIleGlyIleThrPro 20  
DB 151 ATGCGGCTTAAGAGATTGAGCATTCGCAAAAGATGTGCTGATTAATCCCGACAGACA--- 207  
QY 21 LeuValIleuAsnIleuAlaAspGly-CysValAlaArgVal-----AlaAl 37  
DB 208 CTGACCTTTTG-----TTTTTTAGCGAAATTTACAGATCTTACTTAACCTGC 261  
QY 37 alyLeuGluIleuMetGluProCysSerSerValIleAspArgIleGlyIleSerMetIle 57  
DB 262 GAATCTGAATGATGATGAGACCTTCTCAGTTCGAAAGACAGGATAGGTTATGACATGAT 321  
QY 57 eAlaAspAlaGluIleuIleuIleThrProGlyIleSerValIleuIleGluProTh 77  
DB 322 TGTAGATGCTGAGGAGAAAGATTCATCAACCGGAGAAAGTCTCTCATTTGAGCTTAC 381  
QY 77 rSerGlyAsnThrGlyIleGlyIleuAlaPheMetAlaAlaAlaArgGlyTyrIleuIle 97  
DB 382 TAGTGAAGAACCGGAGCATAGTTTGGCATTCAGGAGCTGCTAAGGTTATAAATTAT 441  
QY 97 eIleThrMetProAlaSerMetSerLeuGluArgGlyIleIleuLeuAlaPheGlyAl 117  
DB 442 TATTAACATGCTCTTCGATGAGTCTGAGAGCGAAGCAATCTTCGGGCTTTTGGAGC 501  
QY 117 aGluIleuValIleuThrAspProAlaIleuIleGlyIleValAlaGluIleuIleuIle 137  
DB 502 TGAATTTGCTCTCATCTGATCCCGCAAGGATGAAGAGAGGCTGTTCAAGCAGAGAGA 561  
QY 137 uIleLeuAlaIleuThrProAsnAlaIleIleGluIlePheGluIleuIleuIleuIle 157  
DB 562 GATTCGGAGATGAGCTCAATTAATTTCTTCAATTTGAAATCTTGGCAAGCTTGGCAAGC 621  
QY 157 oIleValIleIleIleIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 177  
DB 622 AAAGATCATTTTGAACCATGAGACCGGAGATCTGAG--GAGCTGATGAGGAGGTTGA 680  
QY 177 palApheValSerGlyIleGlyIleuIleuIleuIleuIleuIleuIleuIleuIleu 197  
DB 681 TGCTCTGCTCTGCGGATGAGCTGAGGATGAGTACGTAACAGGCTGTAAGAAATATCTGAA 740  
QY 197 aGluIleuAsnProAsnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 217  
DB 741 AGAACGAAATCTGATTAAGCTTATGAGCATTTGAACCGATGAAGTCAATATCTGTC 800  
QY 217 rGlyIleuProGlyProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 237

```

DB      801 TGGAGAGAAACCGCGCTCATFAGATCCAGAGAAATGGAGCTTATCCCTGCTGT 860
QY      237 lLeugluValaenLeuLeuAspGluValValGlnIleSerSerAspGluValIleGluTh 257
DB      861 TCTGGAGTGTGATTACTGATTAAGTGTTCATAATTTCAAGTGAAGAGCTATTGGAAC 920
QY      257 rAlaLysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAla 277
DB      921 TGTTAAGCTTCTTGCTGATTAAGAGAGCTTACCTGCGGAATATCATCTGCTGCTGCTG 980
QY      277 aAlaAlaAlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAla 297
DB      981 AGCTGCTGAGTGAATAGATAGCAGAGAGCCAGAAATGCTGGAAGAACTCATTTGCGGT 1040
QY      297 lPheProSerPheGluGluArgGlyLeuSerSerValLeuPheGluSerValArgArgGly 317
DB      1041 CTTCCTCAAGTGTGGAGAGCGCTTATCTATCTGCTGCTGCTTGAATCCGTGAAGCAGA 1100
QY      317 uAlaGluSerMetThrPheGluPro 325
DB      1101 AGCAGAGAGCTGCTGTGAGCCT 1125

RESULT 13
US-10-425-114-22525
/ Sequence 22525, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 22525
/ LENGTH: 1279
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB567-229-H7_F1
US-10-425-114-22525

```

## Alignment Scores:

```

Pred. No.:      5,256-142      Length:      1279
Score:          1309.00      Matches:      256
Percent Similarity: 88.05%      Conservative: 22
Best Local Similarity: 81.13%      Mismatches: 25
Query Match:      80.65%      Indels:      12
DB:              13          Gaps:      1

```

US-09-931-457a-31 (1-325) x US-10-425-114-22525 (1-1279)

```

QY      8 lLeAlaLysAspValThrGluLeuIleGlyLysThrProLeuValTyrLeuAsnLysLeu 27
DB      108 ATGGCCAGAGAGCTCATCCAG-----GTG 131
QY      28 AlaAspGlyCysValAlaArgValAlaAlaLysLeuGluLeuMetGluProCysSerSer 47
DB      132 ACCGAGGGGTGCTCGCGCGCGCTGCGCCAGCTCGAGTCCATGAGCCCTGCTCCAGC 191
QY      48 ValLysAspArgGlyIleGlyTyrSerMetIleAlaAspAlaGluGluLysGlyLeuIleThr 67
DB      192 GTCAAGAGATAGAGATTGGCTTACAGCATGATCAAGAGCCGAGAGAGAGAGGCTGATTACT 251
QY      68 ProGlyLysSerValLeuIleGluProThrSerGlyAsnThrGlyIleGlyLeuAlaPhe 87
DB      252 CCAGGCGTGAAGTGTCTGATTGAAACCACTAGCGGAGACAGGCGCATTTGAGCTGCTTT 311

```

```

QY      88 MetAlaAlaAlaArgGlyTyrLysLeuIleIleThrMetProAlaSerMetSerLeuGlu 107
DB      312 ATGGCTGCTGCCAAGAGCTTACAACTTACACTGCACAAAGCTGCTCATAGAGATGAG 371
QY      108 ArgArgIleIleLeuLeuAlaPheGlyValaGluLeuValLeuThrAspProAlaLysGly 127
DB      372 AGAGAGATCAATATGAGAGCTTTGTGCTTCAACTTGTCTTACTAGACCCATCTCTGGA 431
QY      128 MetLysGlyValaValaGluLysAlaGluGluIleLeuAlaLysThrProAsnAlaTyrIle 147
DB      432 ATGAAAGAGCTGTCAAGAAAGCGAGAGATCAAGAAAGACCCCACTCGTACATC 491
QY      148 LeuGluGlnPheGluAsnProAlaAsnProLysValHisTyrGluThrThrGlyProGlu 167
DB      492 CTTCACAAATTTGAAATCCAGCTAACCCAAAGATTCATATGAGACTACTAGGCTGAA 551
QY      168 lIleTyrLysGlySerAspGlyLysIleAspAlaPheValSerGlyIleGlyThrGlyGly 187
DB      552 ATCTGGAAGCTACAGCAGAAATATCTGCGCTTGTATCTGTGATCGGAGCAGAGAGT 611
QY      188 ThrIleThrGlyAlaGlyLysTyrLeuLysGluGlnAsnProAsnIleLysLeuIleGly 207
DB      612 ACCATCACCGGACTCGGCGAGTACTCAGAGAGCAAAATCTTAATGTCAAGCTCATAGT 671
QY      208 ValGluProValGluSerProValLeuSerGlyGlyLysProGlyProHisLysIleGln 227
DB      672 GTGAGCCAGTGGAGAGTGTGTTTGAATGATGTGAAATCTTGACACACAGATTTCA 731
QY      228 GlyIleGlyAlaGlyPheIleProGlyValLeuGluValaLeuLeuLeuAspGluValaVal 247
DB      732 GGAATGAGAGCTGTTTATCCCTGAGCTTGTGATGTATCTCCCTTGAAGAGCTCTA 791
QY      248 GlnIleSerSerAspGluAlaIleGluThrAlaLysLeuLeuAlaLeuLysGluGlyLeu 267
DB      792 CAGTTTCAAGTGAATGAGATTAAGATGAGCTCCAGAGCTTGTCTTGAAGAGAGGTG 851
QY      268 PheValGlyIleSerSerGlyAlaAlaAlaAlaAlaPheGlnIleAlaLysArgPro 287
DB      852 TTGTTGAGATCTCTTCGTGTCAGCTCGAGCTCGGAGTAAAGCTTGTGAAGAGCCA 911
QY      288 GluAsnAlaGlyLysLeuIleValAlaLysPheProSerPheGlyGluArgTyrLeuSer 307
DB      912 CAAGAGCGCGGAGAGCTATTGTTGTCTGTTCCCGAGCTTCGGAGCGCTTACTCTCA 971
QY      308 SerValLeuPheGluSerValArgArgGluAlaGluSerMetThrPheGluPro 325
DB      972 TGGTGTGCTGTTCCAGTCCATCAAGAGAGCGGAAAGCATGTGTGAGACCC 1025

```

## RESULT 14

```

US-10-425-114-34372
/ Sequence 34372, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 34372
/ LENGTH: 1490
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZWFLMO17256D02_F1
US-10-425-114-34372

```

Alignment Scores:  
 Pred. No.: 2,01e-138 Length: 1490  
 Score: 1279.00 Matches: 245  
 Percent Similarity: 88.38% Conservative: 44  
 Best Local Similarity: 74.92% Mismatches: 36  
 Query Match: 78.80% Indels: 2  
 Gaps: 1

US-09-931-457a-31 (1-325) x US-10-425-114-34372 (1-1490)

Qy 1 MetAlaValGluArgSerGly-----IleAlaIysAspValIrrGluLeuIleGlyLys 18  
 Db 213 GTCCGCCCGAGGTGGGGGCTCAACATCGCAACAGCTCACCCCATCGCGCAAC 272  
 Qy 19 ThrProLeuValIrrLeuAsnLysLeuAlaAspGlyCyValAlaIrrValAlaIlys 38  
 Db 273 ACACCAAGGTGTATCTCAACACGCTCTCAAGGATCTGTGCAATGTGCTGTAG 332  
 Qy 39 LeuGluLeuMetGluProCysSerSerValIysAspArgIleGlyIrrSerMetIleAla 58  
 Db 333 CTCGAGATTAAGAGCCCTGCTGTAGCCTGAAGCAAGATAGGTAAGATATATAT 392  
 Qy 59 AspAlaGluGluValGlyLeuIleThrProGlyLysSerValIleGluProThSer 78  
 Db 393 GATGCTGAACAGAGGGCTGTATCTCTGAAAGAGTGTGTGTGAAGCAACAGT 452  
 Qy 79 GluAsnThrGlyIleGlyLeuAlaPheMetAlaAlaIrrGlyIrrLysLeuIle 98  
 Db 453 GGAACACAGAGCTGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 512  
 Qy 99 ThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValGlu 118  
 Db 513 ACAAGGCTTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 572  
 Qy 119 LeuValIleThrAspProAlaIysGlyMetLysGlyValAlaIrrValAlaGluGlu 138  
 Db 573 CTGTCT 632  
 Qy 139 LeuAlaIysThrProAsnAlaIrrIleLeuGluGluIrrPheGluAsnProAlaAsnProLys 158  
 Db 633 TTAACAGAGACACCAATCTCTCAATGCTTCAACAGTTCATCACTGCCACCTTAAG 692  
 Qy 159 ValIrrGlyIrrThrGlyProGluIleIrrPheGlySerAspGlyLysIleAspAla 178  
 Db 693 GATACCTTATGAGACTGCTGCTCAAGATCTGAGAGATTCACAGGAGAGGTGATATA 752  
 Qy 179 PheValIserGlyIleGlyIrrThrGlyIrrIleThrGlyValAlaGlyLysIrrLeuIle 198  
 Db 753 TTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 812  
 Qy 199 GluAsnProAsnIleLysLeuIleGlyValGluProValGluSerProValIleAsnGly 218  
 Db 813 AAAAATCCCTGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 872  
 Qy 219 GlyLysProGlyIrrProIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrr 238  
 Db 873 GGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932  
 Qy 239 GluValAsnLeuLeuAspGluValAlaIrrIrrSerSerAspGluAlaIrrIrrIrr 258  
 Db 933 GATACCGATATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 992  
 Qy 259 LysLeuLeuAlaIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrr 278  
 Db 993 AACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052  
 Qy 279 AlaAlaPheGluIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrr 298  
 Db 1053 GCCGCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1112  
 Qy 299 ProSerPheGlyIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrrIrr 318

Db 1113 CCGAGCTTCGGCGAGAGTACTTTCATCTGTCTTATCATGCTCATAGAGAAATGC 1172  
 Qy 319 GluSerMetThrPheGluPro 325  
 Db 1173 GAGAACATGAGAGCTGAGCA 1193

US-10-425-114-34527

RESULT 15  
 : Sequence 34527, Application US/10425114  
 : Publication No. US20040034888A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Liu, Jindong  
 : APPLICANT: Zhou, Yihua  
 : APPLICANT: Kovalic, David K.  
 : APPLICANT: Screen, Steven B  
 : APPLICANT: Tabaska, Jack E  
 : APPLICANT: Cao, Yongwei  
 : TITLE OR INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 : FILE REFERENCE: 38-21(53113)B  
 : CURRENT APPLICATION NUMBER: US/10/425,114  
 : NUMBER OF SEQ ID NOS: 73128  
 : SEQ ID NO 34527  
 : LENGTH: 1491  
 : TYPE: DNA  
 : ORGANISM: Zea mays  
 : FEATURe:  
 : OTHER INFORMATION: Clone ID: UC-ZMFLMO17281G01\_FLI  
 : US-10-425-114-34527

Alignment Scores:  
 Pred. No.: 2,01e-138 Length: 1491  
 Score: 1279.00 Matches: 245  
 Percent Similarity: 88.38% Conservative: 44  
 Best Local Similarity: 74.92% Mismatches: 36  
 Query Match: 78.80% Indels: 2  
 Gaps: 1

US-09-931-457a-31 (1-325) x US-10-425-114-34527 (1-1491)

Qy 1 MetAlaValGluArgSerGly-----IleAlaIysAspValIrrGluLeuIleGlyLys 18  
 Db 214 GTCCGCCCGAGGTGGGGGCTCAACATCGCAACAGCTCACCCCATCGCGCAAC 273  
 Qy 19 ThrProLeuValIrrLeuAsnLysLeuAlaAspGlyCyValAlaIrrValAlaIlys 38  
 Db 274 ACACCAAGGTGTATCTCAACACGCTCTCAAGGATCTGTGCAATGTGCTGTAG 333  
 Qy 39 LeuGluLeuMetGluProCysSerSerValIysAspArgIleGlyIrrSerMetIleAla 58  
 Db 334 CTCGAGATTAAGAGCCCTGCTGTAGCCTGAAGCAAGATAGGTAAGATATATAT 393  
 Qy 59 AspAlaGluGluValGlyLeuIleThrProGlyLysSerValIleGluProThSer 78  
 Db 394 GATCCTGAACAGAGGGCTGTATCTCTGAAAGAGTGTGTGTGAAGCAACAGT 453  
 Qy 79 GluAsnThrGlyIleGlyLeuAlaPheMetAlaAlaIrrGlyIrrLysLeuIle 98  
 Db 454 GGAACACAGAGCTGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 513  
 Qy 99 ThrMetProAlaSerMetSerLeuGluArgArgIleIleLeuLeuAlaPheGlyValGlu 118  
 Db 514 ACAATGCTTCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573  
 Qy 119 LeuValIleThrAspProAlaIysGlyMetLysGlyValAlaIrrValAlaGluGlu 138  
 Db 574 CTGTCT 633  
 Qy 139 LeuAlaIysThrProAsnAlaIrrIleLeuGluGluIrrPheGluAsnProAlaAsnProLys 158  
 Db 634 TTAACAGAGACACCAATCTTATCATGCTTCAACAGTTCATCACTGCCACCTTAAG 693

```
QY 159 ValHisTyrGluThrThrGlyProGluHisLeuPylsGlySerAspGlyLysIleAspAla 178
Db 694 GTACATTATAGAGACTACTGCTCCAGAGATCTGGAGAGATTCNAAGGGAGGATGATATA 753
QY 179 PheValSerGlyIleGlyThrGlyGlyThrIleThrGlyAlaGlyLysTyrLeuLysGlu 198
Db 754 TTCAATTGGTGAATTTGAACAGGGGACCAATATCTGTGCCGCCGCTTTTCTCAAGAG 813
QY 199 GluAsnProAsnIleLysLeuIleGlyValGluProValGluSerProValLeuSerGly 218
Db 814 AAAATCCTCGAATTAGGTATTGTAATTGAGCCTTCTGAAAGTAAACATACCTCCGGT 873
QY 219 GlyLysProGlyProHisLysIleGlyValGlyValGlyPheIleProGlyValLeu 238
Db 874 GGAATACTGCTCCACATAGATCCAGGAAATCGCCAGGATTTGTCCAGGAACTTG 933
QY 239 GluValAsnLeuLeuAspGluValValGlnIleSerSerAspGluAlaIleGluThrAla 258
Db 934 GATAGCATATCTTGATGAAGTATGAGATCAAGTATCAAGTATGAGCTGTGACACGCA 993
QY 259 LysLeuLeuAlaLeuLysGluGlyLeuPheValGlyIleSerSerGlyAlaAlaAlaAla 278
Db 994 AAACGTTGCTGCTTCAGAGAGATTACTGCTCGAATCTCCTTGGGGCAGCAGCCGCC 1053
QY 279 AlaAlaPheGlnIleAlaLysArgProGluAsnAlaGlyLysLeuIleValAlaValPhe 298
Db 1054 GCCGCCATTAAGGTGCAAAAGACCAAGAAATCTCGAAATCTCCTTGGGGCAGCAGCCGCC 1113
QY 299 ProSerPheGlyGluArgTyrLeuSerSerValLeuPheGluSerValArgGluAla 318
Db 1114 CCGAGCTTCGGCGAGAGTACCTTCACTGCTTATCAGTCATAGAAGAAATGC 1173
QY 319 GluSerMetThrPheGluPro 325
Db 1174 GAGAACATGACGCTGAGCCA 1194
```

Search completed: May 22, 2004, 01:56:27  
Job time : 588 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 14:15:53 ; Search time 5582 Seconds  
(without alignments)  
10575.638 Million cell updates/sec

Title: US-09-931-457A-30

Perfect score: 1362  
Sequence: 1 acctgtcgtcgcgcagatag.....aaaaaaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

GenEmbl:\*

- 1: gb\_da:\*
- 2: gb\_hcg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pac:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_srs:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_da:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_ov:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_srs:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_hcg\_hum:\*
- 31: em\_hcg\_inv:\*
- 32: em\_hcg\_other:\*
- 33: em\_hcg\_mus:\*
- 34: em\_hcg\_pin:\*
- 35: em\_hcg\_rtd:\*
- 36: em\_hcg\_mam:\*
- 37: em\_hcg\_vrt:\*
- 38: em\_hcg\_vrt:\*
- 39: em\_hcg\_hum:\*
- 40: em\_hcg\_mus:\*
- 41: em\_hcg\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1238	90.9	1267	8	AF452451
2	707.4	51.9	953	8	CAR6024
3	656.4	48.2	1247	8	AY450295
4	645	47.4	1855	8	CNAPCC57
5	643.2	47.2	1354	8	AB029511
6	642.4	47.2	1253	8	AT0ACTLY
7	640.8	47.0	1332	8	ATCY83A
8	638.4	46.7	1280	8	BTY10845
9	636.4	46.7	1281	8	SP10AL
10	636.4	46.7	1303	6	E06716
11	632.8	46.5	1308	8	AF044172
12	632.6	46.4	1233	8	AY063827
13	628	46.1	1254	8	AY045825
14	616	45.4	1238	8	BTY10847
15	609.4	44.7	1192	8	AY353092
16	595.2	43.7	966	6	AX506198
17	595.2	43.7	966	6	AX651825
18	586.4	43.1	1671	8	AB040503
19	557.8	41.0	1378	8	AB029512
20	553.6	40.6	1319	8	ZMCSOATL
21	551.4	40.5	1404	8	AF044173
22	545.2	40.0	993	6	AX652865
23	545.2	40.0	1280	8	AX061993
24	545.2	40.0	1281	8	AK095958
25	543.6	39.9	1290	8	AF073695
26	541.4	39.8	1368	8	ATH011976
27	538.4	39.5	1337	8	AK071279
28	538.4	39.5	1363	8	AF073697
29	530.4	38.9	1355	8	BT009041
30	529.4	38.9	5259	6	AX653890
31	528.8	38.8	1371	8	WHYCY51
32	525.6	38.6	1278	8	ATCYTAC51
33	524.6	38.5	1483	6	E08016
34	524.6	38.5	1484	8	SPICYSY
35	523.2	38.3	1602	8	AK065007
36	522.2	38.3	2104	8	AK065652
37	521.6	38.3	1562	8	AK099217
38	520	38.0	1555	8	AK104704
39	518.2	37.9	1432	8	SOCYGR
40	516.2	37.9	1493	8	NTA29249
41	511.8	37.6	1377	8	ATH271727
42	511.8	37.6	1460	8	AY128885
43	511.8	37.6	1556	8	AY099721
44	509.6	37.4	1210	8	AY096681
45	509.6	37.4	1378	8	AT0ACTLY2

## ALIGNMENTS

RESULT 1  
LOCUS AF452451 1267 bp mRNA linear PLN 21-JAN-2002  
DEFINITION Glycine max cysteine synthase mRNA, complete cds.  
ACCESSION AF452451  
VERSION AF452451.1 GI:18252505  
KEYWORDS  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE 1 (bases 1 to 1267)  
AUTHORS Demoschenko, C. and Krishnan, H.B.

TITLE Nucleotide sequence of a soybean cDNA encoding cysteine synthase  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1267)  
 AUTHORS Demosthenis, C. and Krishnan, H.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-NOV-2001) USDA-ARS, University of Missouri, 108W  
 Curtis Hall, Columbia, MO 65211, USA  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..1267  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 82..1059  
 /codon\_start=1  
 /product="cysteine synthase"  
 /protein\_id="AA66291.1"  
 /db\_xref="GI:18252506"  
 /translation="MAVERSGIADVTLEIGTPLYLANKLADGVAVAKLELMEP  
 CSSTVKRIGYSMTADAEKGLITPKSVLIPTSGNTGIGLMAAARGKLIITMPA  
 SSKSERIILHAFGAEVLTDPKAKMGANVOKABHILATPNAVILQEPENAPRVH  
 YETGPIMKGSDEKIDAVSGITGSTITGAGKILKEQNPILKIVPEVSPVLSG  
 GKQPHKIQGIAGGFIPEGLVNLDBEVIQISSDEALETALKLEGLVGISSGAA  
 AAAAFOIAKRPENNAKLIIVAFPSFGRILSVLPSEVRRESMTPEP"

## CDS

Query Match 90.9%; Score 1238; DB 8; Length 1267;  
 Best Local Similarity 99.6%; Pred. No. 5,9e-291;  
 Matches 1241; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## ORIGIN

Query 20 GCCGATGCTGCTTCTTAGTGTGTCACTCACTTCCTGCTCAAGTCAGCTTTGATG 79  
 Db 12 GCCGATGCTGCTTCTTAGTGTGTCACTCACTTCCTGCTCAAGTCAGCTTTGATG 71  
 QY 80 AGCAGATATATGCTCTTAAAGTCCGGAATTCGCAAGATGTACGGAATTTGATG 139  
 Db 72 AGCAGATATATGCTCTTAAAGTCCGGAATTCGCAAGATGTACGGAATTTGATG 131  
 QY 140 TAAACCCCATTAATATCTTAAATAAATCTGCGATGTGTGTCGCCGCGTGTGTC 199  
 Db 132 TAAACCCCATTAATATCTTAAATAAATCTGCGATGTGTGTCGCCGCGTGTGTC 191  
 QY 200 TAAACTGAGTGTATGAGCCATGCTTATGTCGAGAGACAGAGTTGGTATGAT 259  
 Db 192 TAAACTGAGTGTATGAGCCATGCTTATGTCGAGAGACAGAGTTGGTATGAT 251  
 QY 260 TGCATGATGAGAGAGAGGACTTATCAGACCTGGAAGAGTGTCTCTATTAGGCAAC 319  
 Db 252 TGCATGATGAGAGAGAGGACTTATCAGACCTGGAAGAGTGTCTCTATTAGGCAAC 311  
 QY 320 AAGTGTAAATCTGCGATGTGATTAAGCTTCATGCGACAGCCAGGGGTTACAAGCTCAT 379  
 Db 312 AAGTGTAAATCTGCGATGTGATTAAGCTTCATGCGACAGCCAGGGGTTACAAGCTCAT 371  
 QY 380 AATTACAAATGCTGCTTCTTAGTGTGTGAGAGAGATCATCTATTAGCTTTGGAGC 439  
 Db 372 AATTACAAATGCTGCTTCTTAGTGTGTGAGAGAGATCATCTATTAGCTTTGGAGC 431  
 QY 440 TGAATGTGTTCTGACAGATCCTGCTAAGGAGATGAAGGTGCTTCAGAGGCTGAGA 499  
 Db 432 TGAATGTGTTCTGACAGATCCTGCTAAGGAGATGAAGGTGCTTCAGAGGCTGAGA 491  
 QY 500 GATATGTGCTTAAGAGCCCAATGCTCTATCTTCAATCTTCAATTTGAAAACCTCCCAATCC 559  
 Db 492 GATATGTGCTTAAGAGCCCAATGCTCTATCTTCAATCTTCAATTTGAAAACCTCCCAATCC 551  
 QY 560 CAAAGTTCATTATGAACAACCTGCTCCAGAGATATGGAAGGCTCCAGTGGAAAAATGA 619  
 Db 552 CAAAGTTCATTATGAACAACCTGCTCCAGAGATATGGAAGGCTCCAGTGGAAAAATGA 611  
 QY 620 TGCATTTGTTTCTGAGTACGCACTGCTGTGTACAAATAACAGTCTGGAAAAATATCTTAA 679  
 Db 612 TGCATTTGTTTCTGAGTACGCACTGCTGTGTACAAATAACAGTCTGGAAAAATATCTTAA 671

QY 680 AAGCAGAAATCCGAAATATTAAGCTGATTTGAGAACCACTTGAAGTCCAGTCTCTC 739  
 Db 672 AAGCAGAAATCCGAAATATTAAGCTGATTTGAGAACCACTTGAAGTCCAGTCTCTC 731  
 QY 740 AAGAGAAAGCCGTGTCACACAGATTCAGAGGATTTGGTGTGCTTTATCCCTGATG 799  
 Db 732 AAGAGAAAGCCGTGTCACACAGATTCAGAGGATTTGGTGTGCTTTATCCCTGATG 791  
 QY 800 CTTCGAAATCAATCTTCTTGAATGATTTCAATATCAATGATGAAGCAATGAAAC 859  
 Db 792 CTTCGAAATCAATCTTCTTGAATGATTTCAATATCAATGATGAAGCAATGAAAC 851  
 QY 860 TCGAAAGCTCTTGCGCTTAAAGAGGCTATTTGAGGAAATATCTCCGAGCTGCAGC 919  
 Db 852 TCGAAAGCTCTTGCGCTTAAAGAGGCTATTTGAGGAAATATCTCCGAGCTGCAGC 911  
 QY 920 TGCCTGCTCTTTTCAGATTCGAAAAAGACAGAAAAATGCGGGAAGCTTATTTGCCGT 979  
 Db 912 TGCCTGCTCTTTTCAGATTCGAAAAAGACAGAAAAATGCGGGAAGCTTATTTGCCGT 971  
 QY 980 TTTTCCACCTTCGAGAGAGTACTGTCTCCGTCTATTTGATGTCAGTGAACGCGA 1039  
 Db 972 TTTTCCACCTTCGAGAGAGTACTGTCTCCGTCTATTTGATGTCAGTGAACGCGA 1031  
 QY 1040 AGCTGAAGACATGACTTTTGAAGCCCTGAATTCCTGTTAAGGCTCTCACTGATTTT 1099  
 Db 1032 AGCTGAAGACATGACTTTTGAAGCCCTGAATTCCTGTTAAGGCTCTCACTGATTTT 1091  
 QY 1100 CTGTGTTACTTGTACAGAGCTTTTACTGATTTGTTAGAGTACTACTGTTTGTACCTGAC 1159  
 Db 1092 CTGTGTTACTTGTACAGAGCTTTTACTGATTTGTTAGAGTACTACTGTTTGTACCTGAC 1151  
 QY 1160 TCTAAATAATAACTGCTCCGCAAAAGCTAGTTTCTTCAATGCCCTGAGAGGATATT 1219  
 Db 1152 TCTAAATAATAACTGCTCCGCAAAAGCTAGTTTCTTCAATGCCCTGAGAGGATATT 1211  
 QY 1220 TGTGCTGCAACATTAAAAAGTATTCAAGTGTCTTATTAAGTAAACA 1265  
 Db 1212 TGTGCTGCAACATTAAAAAGTATTCAAGTGTCTTATTAAGTAAAAAAA 1257

## RESULT 2

## LOCUS

CAR6024 953 bp mRNA linear PLN 11-MAY-1998

DEFINITION Cicer arietinum mRNA for cysteine synthase, partial.

VERSION AJ006024.1 GI:3127889

KEYWORDS cysteine synthase

SOURCE Cicer arietinum (chickpea)

ORGANISM Cicer arietinum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae;  
 Cicer.

REFERENCE 1  
 Dopolio, B., Esteban, R. and Labrador, E.  
 cDNA expressed in chickpea epicotyls

JOURNAL Unpublished  
 2 (bases 1 to 953)

REFERENCE  
 AUTHORS Labrador, E.

TITLE Direct Submission

JOURNAL Submitted (04-MAY-1998) Labrador E., Dpto. Fisiologia Vegetal,  
 Univ. Salamanca, Campus Miguel de Unamuno. Pza. Doctores de la  
 Reina s/n, 37007-Salamanca, SPAIN

## FEATURES

## source

1..953  
 /organism="Cicer arietinum"  
 /mol\_type="mRNA"  
 /cultivar="Castellana"  
 /db\_xref="taxon:3827"  
 /clone="Cancys"  
 /tissue\_type="etiolated epicotyls"  
 /dev\_stage="5 day old seedlings"





QY	399	TTGAGCTCTTGAGAGAGAAATGATCTCATTACCTTTTGAGAGCTGAGTGGCTCTGACAGAT	458
Db	304	ATGAGCTCTTGAGAGAGAAATTAATCTCCGCTGCTTTTGCTGAGTGGTACTTAACCAT	363
QY	459	CTGTGTAAGGGAAATGAAAAGTGCTGCTTTCAGAAAGGCTGAAAGATATTTGGCTTAAGACGCC	518
Db	364	CCAGAAAGAGGATCAAAAGGTCTCATCTATCAGAAAGGCTGAAAGAAATTAAGGCCCAACACC	423
QY	519	AATGCTCTACATACCTTCAACAAATTTGAAAACCCCTGCGCATCCCAAGGCTCATTAAGAACC	578
Db	424	AATCTCTTATCTTCACCAATTTGAAAACCCCTGAAAACCAAGAGTACACTATGAGACC	483
QY	579	ACTGGTCAGAGATATGAAAAGGCTCCGATGGGAAATATGATGCAATTTGTTTCGGANTA	638
Db	484	ACTGGTCTGAGATCTGGAAAAGGCTCAATGGGAAAGTAAATGCTCTGCTCTGGAATTT	543
QY	639	GGCACTGTGTGTACATATACAGTGCTGTGAAAATTTCTTAAAGAGCGAATCCCAATATA	698
Db	544	GGAAAGAGAGGCACATATACAGGTTCAAGCAAGTATTTAAGAGAGCAAAACCCGACATA	603
QY	699	AAGCTGATTTGGTGTGAGAACCAAGTTGAAAGTCCAGTGCTCTCAGAGAGAAACCTGTGCCA	758
Db	604	AAGCTCTATGTGTGGAACCAAGTTGAAAGTGTATCTTTCTGAGAGAAACCTGTGCCA	663
QY	759	CACAAAGATTCAAGGATTTGGTGTGCTGCTTTTATCCTGTGTCTTGGAAAGTCAATCTTCTT	818
Db	664	CATAGAGATCCAGGGATATGTGCTGTGTTCAATCTGTGATTTTGGAAAGTGTCTTATTT	723
QY	819	GATGAGTGTGTTCAAAATATCAAGTATGATGAGCAATAGAAACTGCAAAGCTTCTTGCCGCTT	878
Db	724	GATGAGTATTAATTCAGATTTCAAGTATGAGCAATAGAACTGCAAAGCTTCTTGCCGCTG	783
QY	879	AAAGAGGCTCATTTTGTGTGGAATATCTTCCGAGCTGCAAGCTGTGCTGCTTTTCAGATT	938
Db	784	AAAGAGGCTTATGCTTGTGTGGAATATCAATCTGTGCTTCTGCTGCGGCATATCAAACTT	843
QY	939	GCAAAAGAACCAAGAAATATGCCGGAGAGCTTATTTTGCCGATTTTCCAGCTTCGGGAG	998
Db	844	GCTAAGCGGCTTGAAAATGCTGGAAGCTCATTTGTTGTTTTCCTCCAGGCTTCGGGAG	903
QY	999	AGTATCTGTCTCTCCGTCTATTTAGTCAAGTGAAGCGGAAGCTGAAAAGCATGACTTTT	1058
Db	904	CGATATCTTCTCTCTGCTCTTGTGAATCTGTTAAGCGGAAGCAGAGAAACATGACCGTGTG	963
QY	1059	GAGCCCTGAA 1068	
Db	964	GAGCCTTGAA 973	

RESULT 4	1855 bp	mRNA	linear	PLN 28-JAN-2003
CNAPCCS7				
LOCUS				
DESCRIPTION				
ACCESSION	D281777	D17363		
VERSION	0281777.1	GI:366530		
KEYWORDS	O-acetyl-L-serine (thiol)-lyase; O-acetyl-L-serine acetate-lyase (adding hydrogen sulfide); O-acetylserine sulfhydrylase.			
SOURCE	<i>Citruslus lanatus</i> (watermelon)			
ORGANISM	<i>Citruslus lanatus</i>			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eutroids E; Cucurbitales; Cucurbitaceae; Citrullus.			
REFERENCE	1 (bases 1 to 1855)			
AUTHORS	Noji,M., Miyakoshi,I. and Saito,K.			
TITLE	Molecular cloning of a cysteine synthase cDNA from <i>Citruslus lanatus</i> (watermelon) by genetic complementation in an <i>Escherichia coli</i> Cys- auxotroph			
JOURNAL	Mol. Gen. Genet. 244 (1), 57-66 (1994)			
MEDLINE	94316193			
PubMed	8041362			
REFERENCE	2 (bases 1 to 1855)			
AUTHORS	Saito,K.			

**TITLE** Direct Submission  
**JOURNAL** Submitted (28-FEB-1994) *Kazuki Saito, Chiba University, Faculty of Pharmaceutical Sciences, Lab. Mol. Biol. Biotech., Yayocho-1-33, Inage-Ku, Chiba, Chiba 263, Japan (E-mail:ksaito@yayo-u.ac.jp, Tel:81-31-390-2904, Fax:81-43-329-2305)*  
**COMMENT** On Mar 30, 1994 this sequence version replaced g1:391655.  
D17353:Submitted(05-Aug-1993) to by:  
Kazuki Saito, [ksaito@yayo-u.ac.jp](mailto:ksaito@yayo-u.ac.jp)

FEATURES	Location/Qualifiers
source	1. .1855

CDS

[illegible]

## ORIGIN

Query Match	47.48;	Score 645;	DB 8;	Length 1855;
Best Local Similarity	78.68;	Pred. No. 2.1e-146;		
Matches 771;	Conservative	0;	Mismatches 210;	Indels 0;
			Gaps	0;

QY	87	ATAATGCGCTGTGAAGAGTCCGAAATTGCCAAAGATTGTACGGAATTGATTGCTAAACC	146
Db	561	AAATGCGCTGACAAAGTCTACAAATAGCTAAAGATGTTACTGAACATAATGGGAATACCA	620
QY	147	CCATTAGTAATCTAAATTAACCTGGCGAGTGGTGTGTGTCGCCGGGTGCTGCTAAACTG	206
Db	621	CCACTGTAAACCTCAACCGTGTGTGTATGGCTGCGCTCGGGTATGCTGCCAACTTG	680
QY	207	GAGTTGATGAGGACCAATGCTTAGTGTGTGAGGACAGAAATTGGGTATAGTATGATTCGAT	266
Db	661	GAGATGATGAGGCTTGTCTCAAGTGTCAAAATATGGAATTTGGCTATAGCATGATTCAGAT	740
QY	267	GCAAGAGAGAGGACTTATACACCTGAGAAAGATGTCTCTCATTTGAGCCAAAGTGT	326
Db	741	GCAGAAACAGAGGCTTATCACTCCGGGGAGAGTGTCTGAATTGAACCTACAGTGT	800
QY	327	AATATCTGCATTGATTTAGCTTAAATGCGACACCGGGGTTCAAGCTCATATTACA	386
Db	801	AATCTGTGTAAGGTTGGCCCTTCAATGCTGTGCGCAAGGTTATAGCTTATATTGTC	860
QY	387	ATGCCGTCTTATAGAGCTTTGAGAGAAATCAATTCTATTAGCTTTTGAGCTGAGTTG	446
Db	861	ATGCCGTGATTAATAGGCTTTGAAGAAAGAACATTTCTTCAGAGCTTTCGAGCTGAACG	920
QY	447	GTTCTGACAGATCTCTCTAAGGAAATGAAGGTGTCTTCAAGAAAGCTGAAGAGATATG	506
Db	921	GTTCTCAAGATTCAGCTCGGGGAGTGAAGAACACAGTTCAAAAGGCTGAAGAGATTAAG	980
QY	507	GCTAAGAGGCCCAATGCTCATATCTTCAACAATTGAAAAACCTGCGCAATCCCAAGTT	566
Db	981	GCAAAAAACCTTAATCTTACATCTCTTACGAATTTGAAAAACCTGCTAACCCGAAGATC	1046
QY	567	CATTATGAACAACATGCTGTCAAGATATGAAGAAGCTCCGATGGGAAAAATTGATGACTTT	626
Db	1041	CACATGAGACCACTGCTCAGAGATCTGGAGGGGCTCAGGTGGAAGAAATGACACACTTT	1100
QY	627	GTTTCTGGGAATAGGACTGTGTGTACATTAACAGGTCTCGAAAAATATCTTAAAGAGAG	686
Db	1101	GTTCTGTGTAATAGGACTGAGAGTACAGATGACAGGTGACGGGAAGATATCTCAAAAGACA	1166

QY 687 AATCGAATATAAGCTGATGTTGTTGGAACAGTTGAAAGTCCAGTCTCTCAGAGGA 746  
 DB 1161 AATCTAATATTAAGTTATATAGGCTTTGAACTGTTGAAAGCAATTTATCTAGAGGA 1220  
 QY 747 AAGCTGTGTCACACAGATTCAGAGGATTTGCTGTTTATTCCTGTTGTTGAA 806  
 DB 1221 AAGCTGTGTCACACAGATTCAGAGGATTTGCTGTTTATTCCTGTTGTTGAA 1280  
 QY 807 GTCAATCTCTTGAAGAGTTGTTGAAATATCAATGATGAAAGCAATGAAAGTCAAG 866  
 DB 1281 GTCAATCTCTTGAAGAGTTGTTGAAATATCAATGATGAAAGCAATGAAAGTCAAG 1340  
 QY 867 CTTCTTGCGCTTAAGAAAGGCTTATTTGGAATATCTTCGAGCTGCAAGCTGCT 926  
 DB 1341 CTGCTTGCAATGAAAGAGTCTACTGTTGTTATATCATCGGCGCGCGCTGCT 1400  
 QY 927 GCTTTGAGTTGCAAAAAGACCAAGAAATGCGGAGAGCTTATTTGCTTTTCC 986  
 DB 1401 GCAATTTAGGATTCAGAAAGACCAAGAAATGCTGAAAGCTTATTTGCTTTTCC 1460  
 QY 987 AGCTTCGGGAGAGGAGTCTGCTCCGCTATTTGAGTCAAGTGAAGACGAGAGCTGAA 1046  
 DB 1461 AGTTTGGGAGACGATCTCTCAACCGTCTGTTGAGTCTGTGAAAGAGAGACTGAG 1520  
 QY 1047 AGCATGACTTTTGAGCCCTGA 1067  
 DB 1521 AATATGTTTTCAGACCATGA 1541  
 RESULT 5  
 ABO29511 1354 bp mRNA linear PIN 19-JUL-2003  
 LOCUS Solanum tuberosum PCS-1 mRNA for cytosolic cysteine synthase.  
 DEFINITION complete cds.  
 ACCESSION ABO29511.2 GI:32997083  
 VERSION ABO29511  
 KEYWORDS Solanum tuberosum (potato)  
 SOURCE Solanum tuberosum  
 ORGANISM Solanum tuberosum (potato)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 Maruyama, A. and Ishizawa, K.  
 Cytosolic Cysteine Synthase from Potato Sprout  
 JOURNAL Published Only in Database (2001)  
 REFERENCE 2 (bases 1 to 1354)  
 AUTHORS Maruyama, A. and Ishizawa, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUL-1999) Akiko Maruyama, Chiba University, Faculty  
 of Pharmaceutical Sciences, Laboratory of Molecular Biology and  
 Biotechnology, Research Center of Medicinal Resources, Yayoi-cho  
 1-33, Inage-ku, Chiba, Chiba 263-8522, Japan  
 (E-mail:amaru@chiba-u.ac.jp, Tel:81-43-290-2906,  
 Fax:81-43-290-2905)  
 COMMENT On Jul 19, 2003 this sequence version replaced gi:12081916.  
 FEATURES  
 source location/Qualifiers  
 1..1354  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /strain="Dansyaku"  
 /db\_xref="taxon:4113"  
 /cell\_type="sprout"  
 /note="tetraploid"  
 1..1354  
 /gene="PCS-1"  
 85..1062  
 /gene="PCS-1"  
 /codon\_start=1  
 /product="cytosolic cysteine synthase"  
 /protein\_id="BAB20861.1"  
 /db\_xref="GI:12081917"  
 /translation="MAGKIGIGAKDVTLEIGTPIYLVANNVDCVAVARLBSEMP

ORIGIN  
 Query Match 47.2%; Score 643.2; DB 8; Length 1354;  
 Best Local Similarity 76.8%; Pred. No. 6e-146;  
 Matches 786; Conservative 0; Mismatches 228; Indels 0; Gaps 0;  
 CSSVKNRIGYSMTIDAEKRLKPGESVLIEPTSGNTVGGLAFMAAKGYKLITMBS  
 SMLEERILIRAPAEVLVDPKAGMGAISKEIEIAKI PNSTYLDQPEPAPKTH  
 YETTPETIMSGNKGVDALVSGITGGTGLTLEQGNPNVLVYGPESAILSG  
 KRSPHICIGISGAGFIPGVLENVLLIDVVOVSSRESLEMAKLLAKGLVYISSGA  
 AAAIKVAKRPENNAKLIIVIPSPGERIYLSVLEFVREABEMTVBP"  
 47 TCATCTCTGTTCTTCAAGTCAAGCTTTGTAAGTGAAGCAATTAATGCTGTTGAAAGTCT 106  
 DB 42 TCATCTCTGTTCTTTCATTAATCTGAAATTAATGCTGTAATATGCGGGGAAAGAT 101  
 QY 107 CGAAATTCGCAAGATGTAAGCAATGATGTAAGCAATGTAAGCAATGTAAGCAATGTA 166  
 DB 102 TGAATTCGCAAGATGTAAGCAATGATGTAAGCAATGTAAGCAATGTAAGCAATGTA 161  
 QY 167 ACTTCCGAGATGTTGTTGCTCCGAGTTCCTGTAAGCTGATGATGAGCCATGCTC 226  
 DB 162 TGTGTGATGAGGATGTTGCAAGTGTGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCT 221  
 QY 227 TAGGTGAAAGCAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 286  
 DB 222 TAGGTGAAAGCAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 281  
 QY 287 CACACTGGAAGAAGTGTCTCATTTAGCCCAAGAGTGTAAATACGACATTTGATAGC 346  
 DB 282 CAATCCGCGGAGAGTGTCTCATTTAGCCCAAGAGTGTAAATACGACATTTGATAGC 341  
 QY 347 CTTGATGAGCAGACGAGGCTTCAAGCTTATTAATTAATTAATTAATTAATTAATTA 406  
 DB 342 ATTGATGAGCAGACGAGGCTTCAAGCTTATTAATTAATTAATTAATTAATTAATTA 401  
 QY 407 TGAAGAAGATCATTTCTATTAAGCTTTGAGAGTGTGTTGTCAGATCTGCTTAA 466  
 DB 402 TGAAGAAGATCATTTCTGAGGCTTTGAGAGTGTGTTGTCAGATCTGCTTAA 461  
 QY 467 GGAATGAAAGTGTCTGTTCAAGAGCTGAAAGATATTTGGCTTAAGACCCCATGCTA 526  
 DB 462 AGGATGAAAGTGTCTGTTCAAGAGCTGAAAGATATTTGGCTTAAGACCCCATGCTA 521  
 QY 527 CATACCTCAAGATTTGAAACCTGCAATCCCAAGGTTGATTAATTAATTAATTAATTA 586  
 DB 522 TATCTTCAAGATTTGAAACCTGCAATCCCAAGGTTGATTAATTAATTAATTAATTA 581  
 QY 587 AGAGATGAAAGGCTTCGATGGAATTTGATGCAATTTGTTTCTGGATAGGACTGG 646  
 DB 582 TGAATGAAAGGCTTCGATGGAATTTGATGCAATTTGTTTCTGGATAGGACTGG 641  
 QY 647 TGTGAAATTAACAGTGTGTAATATCTTAAGAGAGAGATCGAATTAAGCTGAT 706  
 DB 642 AGGCAATTAACAGTGTGTAATATCTTAAGAGAGAGATCGAATTAAGCTGAT 701  
 QY 707 TGTGTGGAACAGTGTGTAATATCTTAAGAGAGAGATCGAATTAAGCTGAT 766  
 DB 702 TGTGTGGAACAGTGTGTAATATCTTAAGAGAGAGATCGAATTAAGCTGAT 761  
 QY 767 TGAAGGATTTGTTGCTGTTTATCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 826  
 DB 762 TGAAGGATTTGTTGCTGTTTATCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 821  
 QY 827 TGTTCATATCAAGATGTAAGCAATGTAAGCAATGTAAGCAATGTAAGCAATGTAAG 886  
 DB 822 AGTTCATATCAAGATGTAAGCAATGTAAGCAATGTAAGCAATGTAAGCAATGTAAG 881  
 QY 887 CTAATTTGTTGATTAATCTTCGAGAGTGTGCTGCTGCTTTGATGTAAGCAAG 946  
 DB 882 ATTCTGTTGATTAATCTTCGAGAGTGTGCTGCTGCTTTGATGTAAGCAAG 941  
 QY 947 ACCGAAATATCCGGAAGCTTATTTGTTGCGTTTCCAGCTTCGCGGAGAGTACT 1006

Db 942 CCTGAAATGCTGGAGAGCTATGTTGTTATTTCCCAAGCTTGGAGAGGATATCT 1001  
QY 1007 GTCTCCGCTGCTATTTGAGTCACTGAGACGAGAGCTGAAAGCATCTTTGAGCCCTG 1066  
Db 1002 TTCTCTGCTACTCTGCAAACTGTGACAGAGAGAGAGACATGACTGTGAGCCTTG 1061  
QY 1067 AATT 1070  
Db 1062 AACT 1065

RESULT 6  
ATOCACY 1253 bp mRNA linear PLN 15-FBS-2000  
LOCUS Arabidopsis thaliana mRNA for O-acetylserine lyase (At.OAS-5-8).  
DEFINITION X80376.2 GI:6983573  
ACCESSION O-acetylserine(thiol) lyase.  
VERSION Arabidopsis thaliana (thale cress)  
KEYWORDS Arabidopsis thaliana  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta, eudicotyledons, core eudicots;  
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1  
AUTHORS Hell, R., Bork, C., Bogdanova, N., Frolov, I. and Hauschild, R.  
TITLE Isolation and characterization of two cDNAs encoding for  
compartment specific isoforms of O-acetylserine (thiol) lyase from  
Arabidopsis thaliana  
JOURNAL FEBS Lett. 351 (2), 257-262 (1994)  
MEDLINE 94364481  
PUBMED 8082776

REFERENCE 2  
AUTHORS Hell, R.  
TITLE Direct Submission  
JOURNAL Submitted (18-JUL-1994) R. Hell, Ruhr-Universitaet Bochum,  
Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801  
Bochum, FRG

REFERENCE 3  
AUTHORS Hell, R.  
TITLE Revised by (4)  
JOURNAL (bases 1 to 1234)

REFERENCE 4  
AUTHORS Hell, R.  
TITLE Direct Submission  
JOURNAL Submitted (14-FEB-2000) R. Hell, Ruhr-Universitaet Bochum,  
Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801  
Bochum, FRG

REFERENCE 5  
AUTHORS Hell, R.  
TITLE Direct Submission  
JOURNAL Submitted (14-FEB-2000) R. Hell, Ruhr-Universitaet Bochum,  
Lehrstuhl fuer Pflanzenphysiologie, Universitaetstr. 150, 44801  
Bochum, FRG

COMMENT On Feb 16, 2000 this sequence version replaced gi:6032329.

FEATURES  
source location/Qualifiers  
1..1253  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/culivar="Columbia"  
/db\_xref="taxon:3702"  
/clone="At.OAS-5-8"  
/clone\_id="lambda ZAP"  
/dev\_stage="leaf"  
61..1030  
/EC\_number="4.2.99.8"  
/product="O-acetylserine (thiol) lyase"  
/protein\_id="CA56553.2"  
/db\_xref="GI:6983574"  
/db\_xref="GOA:P47998"  
/db\_xref="Swiss-Prot:P47998"  
/translation="MASRIADVTLELGNTPLYLVNNAVCGVRVAALKEKMEPCSS  
VKRIGSMISDAEKGLIKPESVLEPGTGLATPAATGAKGKLIITPASM  
TERRIILAGVLELVLPDPAKMGKAIKAKBEILATPNNYMLQGFENPANIPTIYET  
TGEIWKGTGKIDGTPVSGTCTGTGAKYIKEDNANKYGVPEVSAIIISGGKP  
GPHKIQIGAGTIPSLVNLIDVVOVSSDSIDAKRLAKBGLIVGISGAAAAA

ORIGIN polya\_signal 1153..1159  
AIXLQRPENNAKLFVALFPSPGRISTVLFDATRKBAAMTFEA"  
Query Match 47.2%; Score 642.4; DB 8; Length 1253;  
Best Local Similarity 79.1%; Pred. No. 9.4e-146;  
Matches 763; Conservative 0; Mismatches 201; Indels 0; Gaps 0;  
QY 105 TCCGGAATTCGCAAGATGTTAGCAATGATGTTAAACCCATTGATATCTAAT 164  
Db 68 TCGAATATTCGCAAGATGTTAGCAATGATGTTAAACCCATTGATATCTAAT 127  
QY 165 AAATCTGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 224  
Db 128 AATGTTGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 187  
QY 225 TTTAGTTCGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 284  
Db 188 TTTAGTTCGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 247  
QY 285 ATCAACCTGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 344  
Db 248 ATCAACCTGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 307  
QY 345 GCTTCATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 404  
Db 308 GATTCATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 367  
QY 405 CTGAGAGAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 464  
Db 368 ACTGAGAGAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 427  
QY 465 AAGGAGATGAAAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 524  
Db 428 AAGGAGATGAAAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 487  
QY 525 TACATGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 584  
Db 488 TACATGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 547  
QY 585 CCAGATGATGAAAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 644  
Db 548 CCAGATGATGAAAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 607  
QY 645 GGTGATGCAATGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 704  
Db 608 GGTGATGCAATGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 667  
QY 705 ATGATGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 764  
Db 668 ATGATGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 727  
QY 765 ATTCAGAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 824  
Db 728 ATTCAGAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 787  
QY 825 GTTGTGCAATGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 884  
Db 788 GTTGTGCAATGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 847  
QY 885 GGTGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 944  
Db 848 GGTGATGCAAGATGTTAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 907  
QY 945 AAGCAGGAAATGCGCGAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 1004  
Db 908 AAGCAGGAAATGCGCGAGTTCGCGAGTTCGCAATGATGTTAAACCCATTGATATCTAAT 967  
QY 1005 CTGCTCTCGTCTATTTAGTCACTGAGACGAGAGCTGAAGAGATGATCTTTAGGCC 1064  
Db 968 CTATCAAGGTACTTTGATGAGCAAGAGAAAGAGCGAGAGCATGACTTTCAGGCT 1027  
QY 1065 TGA 1068



L.: evidence for Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform  
 JOURNAL Plant Mol. Biol. 37 (1), 87-97 (1998)  
 MEDLINE 98281577  
 PUBMED 9620267  
 AUTHORS 2 (bases 1 to 1280)  
 DIRECT SUBMISSION Schaefer, H.J.  
 TITLE Submitted (28-JAN-1997) H.J. Schaefer, Universitaet Heidelberg,  
 JOURNAL Botanisches Institut, Im Neuenheimer Feld 360, 69120 Heidelberg,  
 FRG

# FEATURES

source  
 1. 1280  
 /organism="Brassica juncea"  
 /mol\_type="mRNA"  
 /sub\_species="Viltasso"  
 /db\_xref="taxon:3707"  
 /clone="OAS-T14"  
 /issue\_type="root"  
 /clone\_lib="lambda ZAP Express"  
 /dev\_stage="mature"  
 74. 1042  
 /codon\_start=1  
 /product="O-acetylserine(thiol) lyase"  
 /protein\_id="CAA71798.1"  
 /db\_xref="GI:2243120"  
 /db\_xref="GOA:O23733"  
 /db\_xref="SWISS-PROT:O23733"  
 /translation="MASRLKDYTHLIGTPVLYLNNVAGCGVRAALFEMRPS  
 VKRIGSMISDAEQKGLKPSVILIPSGATGGLAFLPAARVILITPMSK  
 VERRIILAFAGVLLTPAKMKALIAAEELATPEVYMLQFENPKNHET  
 TGEPIWKGDKIPFVSGIGTGITIGAKLMEONPVKLYGVEPIISALISGKE  
 GPRIQIGAGFIPVLEVDLIDEVQVSDSIDMARLLALKEGHLVIGISSGAAAA  
 ATLAKRPMAGLFLVAVPFSFERILSTVLPATRKESMTFQA"

Query Match 46.7%; Score 636.4; DB 8; Length 1280;  
 Best Local Similarity 78.7%; Pred. No. 2,7e-144;  
 Matches 760; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

## ORIGIN

105 TCCGGAATGGCAAGATGTTACGGAATGATGTTAAACCCCATAGTATCTAAAT 164  
 80 TCTGAAATGCTAAAGATGTTACGGAATGATGTTAAACCCCATAGTATCTAAAT 139  
 165 AAACCTGCGATGTTGTTGCGGCTGCTGCTAACTGATGATGATGATGATGATG 224  
 140 AATGCTGCTAAAGATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199  
 225 TCTAGCTGAAAGACGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 254  
 200 TCCGATGCTAAAGACGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259  
 285 ATCAACCTGAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344  
 260 ATCAACCTGAAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319  
 345 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404  
 320 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379  
 405 CTGAGAGAGAGATCATTTCTTATGCTTTGAGCTGAGTGGTGTGACAGATCTGCT 464  
 380 GTGAGAGAGAGATCATTTCTTATGCTTTGAGCTGAGTGGTGTGACAGATCTGCT 439  
 465 AAGGAGATGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524  
 440 AAGGAGATGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499  
 525 TACATCTTCAACATTTGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584  
 500 TACATCTTCAACATTTGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559  
 585 CCAAGATATGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644

560 CTTGAGATATGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619  
 645 GGTGATCAATTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704  
 620 GGTGATCAATTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679  
 705 ATTGATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764  
 680 TATGAGATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739  
 765 ATTCAAGGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824  
 740 ATTCAAGGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799  
 825 GTTGTCAATTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884  
 800 GTTGTCAATTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859  
 885 GGCCTATTTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944  
 860 GGTCTCTTGTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919  
 945 AACCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004  
 920 AACCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979  
 1005 CTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064  
 980 CTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039  
 1065 TGAAT 1070  
 1040 TGAAT 1045

## RESULT 9

SPOTIAL 1281 bp mRNA linear PIN 02-FEB-1999  
 LOCUS Spinach mRNA for O-acetylserine(thiol)-lyase, complete cds.  
 DEFINITION D10476  
 ACCESSION D10476.1 GI:218276  
 VERSION O-acetyl-L-serine acetate-lyase (adding hydrogen sulfide);  
 KEYWORDS O-acetylserine(thiol) lyase; cysteine synthase.  
 SOURCE Spinacia oleracea (spinach)  
 ORGANISM Spinacia oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Caryophyllales; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Amaranthaceae; Spinacia.  
 REFERENCE 1 (bases 1 to 1281)  
 AUTHORS Saito, K., Miura, N., Yamazaki, M., Hirano, H. and Murekoshi, I.  
 TITLE Molecular cloning and bacterial expression of cDNA encoding a plant  
 cysteine synthase  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8078-8082 (1992)  
 MEDLINE 92390392  
 PUBMED 1518633  
 2 (bases 1 to 1281)  
 Saito, K.  
 DIRECT SUBMISSION  
 TITLE Submitted (12-FEB-1992) Kazuki Saito, Chiba University, Faculty of  
 Pharmaceutical Sciences, Lab. Mol. Biol. Biotech., Yayoi-cho 1-33,  
 Inage-ku, Chiba, Chiba 263, Japan (E-mail:ksaito@ph.chiba-u.ac.jp,  
 Tel:043-290-2904, Fax:043-290-2905)  
 COMMENT Submitted (12-FEB-1992) to DDBJ by:  
 Faculty of Pharmaceutical Sciences,  
 Chiba University  
 Yayoi-cho 1-33  
 Chiba 260  
 Japan  
 Phone: 0472-51-1111 x2707  
 Fax: 0472-55-1574.  
 Location/Qualifiers  
 1. 1281





169 GAGGAGTGAACCTTCTCTAGTGTGAAGACAGATGGGTTGCTGAGTATGATCTGAT 228  
 267 GCAGAGAGAGAGAGACTTATACACCTGGAAGAGTGTCTCTCATGAGCCACAGAGTGT 326  
 229 GCTGAAAAAGCGGCTTATACACCTGAGAGAGTGTCTGATGAGCCACAGAGTGTGA 288  
 327 AATACGCGCATGATTAAGCTTCATGCGAGAGCCAGGGGTTTCAAGCTCATATATGA 386  
 289 AATACGCGCATGATTAAGCTTCATGCGAGAGCTTAAGGTTTCAAGCTCATATATGA 348  
 387 ATGCTGCTCTCTATGATGCTTGAAGAGATATCATTTGATGCTTGTGAGCTGAGTTG 446  
 349 ATGCGAGCATCATGATGCTTGAAGAGATATCTCAGGGCTTGTGCTGAGCTT 408  
 447 GTTCTGACATCTCTGATGAGGATTAAGAGTGTCTTGAAGGCTGAGAGATATG 506  
 409 ATCTTACTGATCCAGAGAGAGATGAGAGGCTGTGTGAGAGGCTGAGAGATCCGT 468  
 507 GCTAAGAGCGCCATGCTCATGATCTTCAATTTGAAAAACCTGCGCAATCCCAAGTT 566  
 469 GACAAACTCTCTATGATTAATATATATATATATATATATATATATATATATATAT 528  
 567 CATTTGAAACCACTGCTCCAGAGATATGAGAAAGCTCCATGAGGAAATTTGATGAT 626  
 529 CATTTGAAACCACTGAGCCAGAAATTTGAGAAAGGCAAGGTGAGAAATTTGATGAT 588  
 627 GTTCTGAGATGAGCACTGAGTGTATATATATATATATATATATATATATATATAT 686  
 589 GTTCTGAGATGAGCACTGAGTGTATATATATATATATATATATATATATATATAT 648  
 687 AATCCGAT 746  
 649 AATCCGAT 708  
 747 AATCCGAT 806  
 709 AATCCGAT 768  
 807 GTCAATCTTTTGAAGATTTGCAATATATATATATATATATATATATATATATAT 866  
 769 GTCAATCTTTTGAAGATTTGCAATATATATATATATATATATATATATATATAT 828  
 867 CTCTTGGGCTTAAGAGGCTATTTGAGGAAATATCTTCCGAGGCTGAGCTGCT 926  
 829 TTGCTGCTCTCAAGAGAGGCTTCTGCTGAGGATTTCTGCTGCTGCTGCTGCT 888  
 927 GCTTTGAGATTTGAGAGAGGCTTCTGCTGAGGATTTCTGCTGCTGCTGCTGCT 986  
 889 GCTTTGAGATTTGAGAGAGGCTTCTGCTGAGGATTTCTGCTGCTGCTGCTGCT 948  
 987 AGCTTCGGGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046  
 949 AGCTTCGGGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008  
 1047 AGCATGACTTTTGAAGGCTGATTTCCGTT 1076  
 1009 AGCATGACTTTTGAAGGCTGATTTCCGTT 1038

RESULT 11  
 AF044172 1308 bp mRNA linear PLN 06-JUL-1998  
 LOCUS Solanum tuberosum cysteine synthase mRNA, complete cds.  
 DEFINITION Solanum tuberosum cysteine synthase mRNA, complete cds.  
 ACCESSION AF044172  
 VERSION AF044172.1 GI:3290019

KEYWORDS Solanum tuberosum (potato)  
 SOURCE Solanum tuberosum  
 ORGANISM Eubacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 1308)

AUTHORS Hesse, H. and Hoefgen, R.  
 TITLE Isolation of cDNAs encoding cytosolic (Accession No. AF044172) and  
 plasmidic (Accession No. AF044173) cysteine synthase isoforms from  
 Solanum tuberosum (PGR98-057)  
 JOURNAL Plant Physiol. 116, 1604 (1998)  
 REFERENCE 2 (bases 1 to 1308)  
 AUTHORS Hesse, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-JAN-1998) Freie Universitaet Berlin, Institut fuer  
 Angewandte Genetik, Albrecht-Flaer-Weg 6, Berlin 14195, Germany

## FEATURES

source  
 1..1308  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Berolina"  
 /db\_xref="taxon:4113"  
 35..1012  
 /note="CS-A: O-acetylserine (thiol) lyase; cytosolic  
 isoform"  
 /codon\_start=1  
 /product="cysteine synthase"  
 /protein\_id="AAC25635.1"  
 /db\_xref="gi:3290020"  
 /translation="MAGEKIGIAKDVTEIIGNTPLVYINNVYDGCARVAMLESMER  
 CSVDRDIGSMITDPERKGLKPGESVLIETSGNTGVLAFMAAGYKLIITWPS  
 SMELRRITIRGPRSEIVITDPAKMKAKISAEIRIKATPNSYTILOEPENPAPKIH  
 YHTPEIKTNGSKVDALASGIGTGCTTSGRYLPRQNPVYKIGVPEVPSATISG  
 GPGPHKIQGIAGFTPGVLEYNLIDVAVVSSSESIEMAKILAKELIVGISGGA  
 AAAIKVAPRPNMAGLIVLPSPSGERYLSVLPETVRRKENVTEP"

## ORIGIN

Query Match 46.5%; Score 632.8; DB 8; Length 1308;  
 Best Local Similarity 76.9%; Pred. No. 26-143;  
 Matches 772; Conservative 0; Mismatches 232; Indels 0; Gaps 0;  
 67 AAGCTTTGAGAGAGATTAATGCTGTTGAAGGCTCCGAAATTTGCCAAGATGTA 126  
 12 AATCTGAGATTAATGCTGATTAATGCGCGGAGAAAGATGGAATGCGAAGATGA 71  
 127 CGGAATGATGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 186  
 72 CGGAATGATGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 131  
 187 CCGGAGTCTGCTAACTGAGATTAATGCTGATTAATGCTGATTAATGCTGATTA 246  
 132 CAGGTGTCTGCTGAGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTG 191  
 247 GGTATGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTG 306  
 192 GTTATGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTG 251  
 307 TCATTGAGCAACAAGTGTAACTGAGATTAATGCTGATTAATGCTGATTAATGCTG 366  
 252 TCAATGAGCAACAAGTGTAACTGAGATTAATGCTGATTAATGCTGATTAATGCTG 311  
 367 GTTACAGCTCATATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 426  
 312 GCTTACAGCTCATATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 371  
 427 TAGCTTTGAGAGTGTGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTAATG 486  
 372 GTGCTTTGAGAGTGTGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTAATG 431  
 487 AGAAGCTGAGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 546  
 432 CAAGGCTGAGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGATTA 491  
 547 ACCCTGCAATCCCAAGTTCATTAATGAGCACTGCTCAGAGATTAATGAGAGGCTCC 606  
 492 ACCCTGCAATCCCAAGTTCATTAATGAGCACTGCTCAGAGATTAATGAGAGGCTCA 551  
 607 ATGGAGAAATTAATGATTTGTTTCTGAGATTAATGAGCACTGCTCAGAGATTAATG 666



Db 552 ACGGGAAGTGAAGTCTAGCTCTGATTTGGAACAGAGGACGATTAAGTGTTCAG 611

Qy 667 GAAATATCTTAAAGAGAGAAATCCGAATATTAAGCTGATTTGGTGAACCAAGTTGAA 726

Db 612 GCAAGTATTTGAGAGAGAGAACCCCAACGTTAGCTTATGCGCTGGAACAGATTGAA 671

Qy 727 GTCCAGTCTCTCAAGAGAGAAAGCTGTGTCACACAGATTTCAAGGATTTGGTGTGTT 786

Db 672 GTGCTATCTTTCTGTGTGGAAGCTTGTGTCACATTAAGTTCAAGGATTTGGTGTGTT 731

Qy 787 TTATCCCTGATCTCTGTAAGTCAATCTTTGATGAAGTTGTTCAATATCAAGTATG 846

Db 732 TCATCTCTGATGTTTGAAGTAACTTATGATGATGATGATGATGATGATGATGATG 791

Qy 847 AAGCAATGAAGTGAAGTCTTCTTGGCTTAAGAGGCTTATTTGGTGAATATCTT 906

Db 792 AATCATGAAGTGAAGTCTTCTTGGCTTAAGAGGCTTATTTGGTGAATATCTT 851

Qy 907 CCGAGAGTGAAGTCTCTGCTCTTCAATTTGCAAAAAGACAGAAAATGCGGAGAGC 966

Db 852 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911

Qy 967 TTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026

Db 912 TCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971

Qy 1027 CAGTGAAG 1070

Db 972 CTGTCAAG 971

RESULT 12  
LOCUS AY063827  
DEFINITION Arabidopsis thaliana At3g22460 mRNA sequence.  
ACCESSION AY063827.1 GI:17380705  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses; 1 (bases 1 to 1233)

REFERENCE  
AUTHORS Yamada,K., Banu,J., Banno,F., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,U., Kim,C.J., Narisaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
Arabidopsis Full Length cDNA Clones  
Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS 2 (bases 1 to 1233)  
Yamada,K., Banu,J., Banno,F., Chang,C.H., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,U., Kim,C.J., Narisaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
Direct Submission  
Submitted (13-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

TITLE  
JOURNAL  
REFERENCE  
AUTHORS 3 (bases 1 to 1233)  
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,U., Kim,C.J., Narisaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (23-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNA (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'); Seki,M., Narisaka,M., Ishida,J., Hayashizaki,Y. and Shinozaki,K.

FEATURES  
SOURCE  
1..1233  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAFL06-87-B19 (R11881)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SacI insert."

gene  
misc\_feature  
1..1233  
/gene="At3g22460"  
84..1061  
/gene="At3g22460"  
/note="This may be a potential product of a partially processed transcript corresponding to gene At3g22460."  
/evidence="experimental"

misc\_difference  
1216  
/gene="At3g22460"  
/note="not present in genomic sequence"

ORIGIN  
Query Match 46.4%; Score 632.6; DB 8; Length 1233;  
Best Local Similarity 77.1%; Pred. No.2.3e-143;  
Matches 770; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 71 TTGTAGTGAAGCAATTAATGCTGTGTAAGGTCGGAATGCCAAAGTGTACGA 130

Db 65 TTTTCAAGTGAAGTGAACATGCTTCAAGTGTCTTAAATGCTTAAAGATTAACGA 124

Qy 131 ATTGATGTGAAGAACCCATTAAGTATCTTAATTAACCTTGGATGTGTGTCGCG 190

Db 125 ACTAATGGAAGACCTCATTAGTATCTTAACCAAGTGAAGCTTAAGATTTGTGTC 184

Qy 191 GATTGCTGTAACTGAGTGTGATGAGGCAATGCTAGTGTGAAGACAGATTTGCTA 250

Db 185 TGTTCGCTAGAGTTGAATGATGAGAACCATGTGTAGAGTCAAAAGACAGATTTG 244

Qy 251 TAGTATATGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310

Db 245 TAGTATATGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304

Qy 311 TGAGCCAAAGAGTATATCTGAGTGTGATGAGGCAATGCTTATGAGGACAGAGGCTTA 370

Db 305 TGAGCCAAAGAGTATATCTGAGTGTGATGAGGCAATGCTTATGAGGACAGAGGCTTA 364

Qy 371 CAAGCTATATTAATCAATGCTGCTTATGAGTCTTGAAGAGAGAGATCTATTAAGC 430

Db 365 TTAGCTTGTATTAACAATGCACTTCAAGATTAATGAAGAGAGAGATCAATCTTATAGC 424

Qy 431 TTTTGAAGTGAAGTGTGTTTGAAGATCTGCTGAAGAGAGAGAGAGAGAGAGAGAG 490

Db 425 TTTTGAAGTGAAGTGTATTTTATGATGATGATGATGATGATGATGATGATGATG 484

QY 491 GGCTGAGAGATATGGCTAAGAGCCCAATGCTACTTCAACATTTGAAAAACC 550  
 DB 485 GGGGAGAGATTTTGGCCAAAACCTATATGTTATCTTCAACAGTTTGAAGATCC 544  
 QY 551 TGGCAATCCCAAGCTTATTATGAACCACTGCTCCAGATATGAAAGGCTCCATG 610  
 DB 545 TGGTAATCCAAAGATCATTAATGAGCTACTGAGCCGAGATATGAAAGGCTCTG 604  
 QY 611 GAAATATGATGCAATTTGTTCTGGAGTAGGCACTGGTGTATCAATPAAACAGTCTG 670  
 DB 605 AAAAGTGAATGCTTTTCTGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 664  
 QY 671 ATATCTTAAAGAGAGATCCGAATATTAAGCTGATGTTGTTGAAACAGTTGAAAGTCC 730  
 DB 665 GTATCTCAAGAGACAGAAACAAACATTAAGCTGATGTTGTTGAAAGCTTGAAGCC 724  
 QY 731 AGTCTCTGAGAGAAAGGCTGCTCCACAAATGATCAGAGGATGTTGCTGTTTAT 790  
 DB 725 TATTTCTGCTGGAGAGAGGCTGCTCCACAAATGATGTTGTTGTTAT 784  
 QY 791 CCTGCTGCTGAGAGATCATCTTTGATGAGTGTCAATATCAAGTATGATGAGC 850  
 DB 785 TCCAGGATTTTGGATGTTATCTTAATGATGAGTGTCAAGTATGATGAGC 844  
 QY 851 AATTAAGAACTGCAAGCTTCTTGGCTTAAGAGGCTTATTTGTTGTTATCTTCCG 910  
 DB 845 CATTAAGTCAAGAGGCTTCTTGGAGAGAGGCTTATTTGTTGTTATCTTCCG 904  
 QY 911 AGCTGAGAGCTGCTGCTTCTTTGATGAGTGTCAAAAGACAGAAATGCGGAGAGCTTAT 970  
 DB 905 AGGAGAGAGCTGCTGCTGAGCATCAAACTGAAAGAGGCTTATTTGTTGTTATCTTCCG 964  
 QY 971 TGTTCCTGTTTCCAGCTTCCGAGAGAGTCTCTCTGCTGTTATGATGATG 1030  
 DB 965 AGTGGCGGTGTTCCGAGTTCGGGAGAGATATCTTAATGATGATGATG 1024  
 QY 1031 GAGAGCGAGAGTGAAGCATGCTTTGAGCCCTGAT 1069  
 DB 1025 AAGGAAAGAGCGAGAGCTATGACCTTTGAGCCCTGAT 1063

RESULT 13  
 AY045825  
 LOCUS  
 DEFINITION Arabidopsis thaliana putative cytosolic O-acetylserine(thiol) lyase (At4g14880) mRNA, complete cds.  
 ACCESSION AY045825  
 VERSION AY045825.1  
 KEYWORDS GI:15027936  
 SOURCE F1 CDNA.  
 ORGANISM Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1254)  
 Yamaoka, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Full Length cDNA Clones  
 Unpublished  
 2 (bases 1 to 1254)  
 Yamaoka, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, B., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tracy, S.B.,

TITLE  
JOURNAL  
COMMENT

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Direct Submission  
 Submitted (10-JUN-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-Length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.)

The Sak, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamaoka, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, B., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tracy, S.B., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamaoka, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES  
source

1..1254  
 location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 /clone="RAF05-20-P15 (R10445)"  
 /ecotype="Columbia"  
 /note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SacI insert."  
 1..1254  
 /gene="At4g14880"  
 1..77  
 /gene="At4g14880"  
 78..1043  
 /gene="At4g14880"  
 /note="contains cysteine synthase/cystathionine beta-synthase P-phosphate attachment site A35-53"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="putative cytosolic O-acetylserine(thiol) lyase"  
 /protein\_id="AAK76499.1"  
 /db\_xref="GI:15027937"  
 /translation="MRSRIADYVETLIGNTPVANNVAGCGVRAATLBMPECCSVKDRIGFNSIDMKKGLIKPGESVITETSGNTGATGAPTAAPAKKYLITMPASMTERRITLAFGVELVTDPAKAKKAIAPAEELATPAGTAKYGVSPVSAALISGKPGPEIKWKGKIDGFSVGIIGGTTGAKXKLEKANNVAKYGVSPVSAALISGKPGPHKIQIGAGPSTVANDLIDDEVVQSDSIDMARQLAKEGHLLVQISSGAAAAA IKLAQRPNAGLFPVAFIPSPFGERYLTVLEPATKXKAEAMTFEA"

misc\_difference  
 /gene="At4g14880"  
 /note="compared to genomic sequence"  
 /replace="gtaa"  
 misc\_difference  
 /gene="At4g14880"  
 /note="compared to genomic sequence"  
 /replace="tttg"  
 3'UTR  
 /gene="At4g14880"  
 /replace="tttg"  
 misc\_difference  
 /gene="At4g14880"  
 /note="not present in genomic sequence"

## ORIGIN

Query Match 46.1%; Score 628; DB 8; Length 1254;  
 Best Local Similarity 78.9%; Pred. No. 3e-142;  
 Matches 761; Conservative 0; Mismatches 200; Indels 3; Gaps 1;



QY 408 GAGAGAAATCATCTTATAGCTTTTGGAGCTGAGTGGTCTGACAGATCCTGCTAAG 467  
 Db 333 GAGAGAAATCATCTTATAGCTTTTGGAGCTGAGTGGTCTGACAGATCCTGCTAAG 392  
 QY 468 GGAATGAAGTGTCTTCAAGAGCTGAGAGATTTGGCTAGAGCCCAATGCTTAC 527  
 Db 393 GGCATGAAGGCTATATCGCAAGAGCGAGAGATTTGGCTAGAGCCCAATGCTTAC 452  
 QY 528 ATACTTCACATTTTGAAGAACCTTCGCAATCCCAAGCTTCATTAAGAACCACTGCTCA 587  
 Db 453 ATGCTCCACAGCTTTGAGAACCTTCGCAATCCCAAGCTTCATTAAGAACCACTGCTCA 512  
 QY 588 GAGATGAAGAGCTCGATGAGAGAAATTAATGATTTGTTTGGATAGGACCTGCT 647  
 Db 513 GAGATTTGAGAGGACATGAGAGGATTAATGATTTGTTTGGATAGGACCTGCT 572  
 QY 648 GGTACATTAACAGGTGCTGAGAAATATCTTAAGAGAGATCCGAATATTAAGCTGAT 707  
 Db 573 GGTACATTAACAGGTGCTGAGAGATATCTTAAGAGAGATCCGAATATTAAGCTGAT 632  
 QY 708 GGTGTGACACGATTTGAAGATTCAGTGTCTTCAAGAGAGAGCTTGTCCACAGAT 767  
 Db 633 GGTGTGACACGATTTGAAGATTCAGTGTCTTCAAGAGAGAGCTTGTCCACAGAT 692  
 QY 768 CAAGGATTTGCTGCTGTTTATCCCTGCTGCTTGAAGATCAATCTTCTGATGAGT 827  
 Db 693 CAAGGATTTGCTGCTGTTTATCCCTGCTGCTTGAAGATCAATCTTCTGATGAGT 752  
 QY 828 GTTCAATATCAAGTATGATGAAGCAATGAAGAACTGCTTCTGCTTGAAGAGC 887  
 Db 753 GTTCAATATCAAGTATGATGAAGCAATGAAGAACTGCTTCTGCTTGAAGAGC 812  
 QY 888 CTATTGTTGGAATATCTTCCGAGCTGAGCTGCTGCTTCTGATTTGCAAGAA 947  
 Db 813 CTATTGTTGGAATATCTTCCGAGCTGAGCTGCTGCTTCTGATTTGCAAGAA 872  
 QY 948 CCAGAAATGCGGAGAGCTTATTTGCTGCTTCTGCTTCCAGCTTCCGAGAGTACTG 1007  
 Db 873 CCAGAAATGCGGAGAGCTTATTTGCTGCTTCTGCTTCCAGCTTCCGAGAGTACTG 932  
 QY 1008 TCTTCCCTGCTTATTTGATCTGATGAGAGCGAGAGCTTGAAGATCTTTGAGACCTG 1067  
 Db 933 TCGACCTGCTTCTGATGAGCAAGAGAGAGAGATCCATGATCTTGAAGCTTGA 992  
 QY 1068 AT 1069  
 Db 993 AT 994

RESULT 15  
 AY353092 1192 bp mRNA linear PLAN 27-AUG-2003  
 LOCUS Populus x canescens O-acetylserine (thiol) lyase mRNA, partial cds.  
 DEFINITION AY353092  
 ACCESSION AY353092.1 GI:34099832  
 KEYWORDS  
 ORGANISM  
 SOURCE  
 Populus alba x Populus tremula  
 Populus alba x Populus tremula  
 Burkholderia, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
 1 (bases 1 to 1192)  
 Kopriva, S., Hartmann, T., Messaro, G., Hoenicke, P. and Rennerberg, H.  
 Regulation of sulfur assimilation by nitrogen and sulfur nutrition  
 in poplar trees  
 Unpublished  
 2 (bases 1 to 1192)  
 Kopriva, S.  
 Direct Submission  
 Submitted (29-JUL-2003) Institute of Tree Physiology,  
 Georg-August-Universität Göttingen, Göttingen, Germany  
 Location/Qualifiers  
 1..1192  
 source

ORIGIN  
 Query Match 44.7%; Score 609.4; DB 8; Length 1192;  
 Best Local Similarity 79.9%; Pred. No. 1e-137;  
 Matches 718; Conservative 0; Mismatches 181; Indels 0; Gaps 0;  
 CDS  
 /organism="Populus alba x Populus tremula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:80863"  
 /clone="INRA 717 1B4"  
 <1..879  
 /note="putative cytosolic isoform"  
 /codon\_start=1  
 /product="O-acetylserine (thiol) lyase"  
 /protein\_id="AA057205.1"  
 /db\_xref="GI:34099832"  
 /translation="RIAALEEMEPDPSVYDRIGYSMTIDAEKGLIKAGEVLEPT  
 SCNVTGLAFMAAGYRLITMPASMSLERPVLLAFGAELVLEPAAGMGAQVKA  
 ERISAKTPSYLLOEPENANPKVHTDTPPIHSGSGGKQAPVSGDITGGITGAG  
 KYLKEONPKIKVGHPEVNSAVLSGSKRPKICIGAGPIGIVNVDLIDETVVIS  
 EALIFAKLALKEGLVGISGAAAMAIKAKPEPNAGKLIVAFPSFGRIYLSV  
 LFEVYKKEENWFEF"  
 189 CCGGTTGCTGCTAACTGAGATTGATGAGACCACTCTGATGTGAAGACAGATTGGG 248  
 1 CGAATGCTGCGCAAGCTGAGATGATGAACTTGTGAGTGTCAAGACAGATTGGG 60  
 249 TATAGTATGATTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 308  
 61 TATAGTATGATTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 309 ATTAGCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 368  
 121 ATTAGCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 369 TACAGCTCATTTATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428  
 181 TACAGCTCATTTATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 429 GCTTTGAGCTGATGTTGTTCTGACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 488  
 241 GCTTTGAGCTGATGTTGTTCTGACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 489 AAGCTGGAAGATTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548  
 301 AAGCTGGAAGATTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 549 CTGCGCAATCCCAAGTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608  
 361 CTGCGCAATCCCAAGTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 609 GGGAAATGATGATTTGTTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668  
 421 GGGAAATGATGATTTGTTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 669 AAATATCTTAAG 728  
 481 AATATCTTAAG 540  
 729 CAGTGTCTGAG 788  
 541 CAGTGTCTGAG 600  
 789 ATCCCTGCTGCTGAG 848  
 601 ATCCCTGCTGCTGAG 660  
 849 GCAATGAG 908  
 661 GCAATGAG 720  
 909 GAGCTGAG 968  
 721 GAGCTGAG 780

Mon May 24 08:18:49 2004

us-09-931-457a-30.rge

Page 15

Qy 969 ATTGTTGCCGTTTTCCTCCAGCTTCGGGAGAGGTACTGTCTCCGTCCTATTGAGTCA 1028  
Db 781 ATTGTTGCCGATCTTCCCTAGTTTGGAGAGCGTTACTATCCTCTGTGCTTTTGAATCA 840  
Qy 1029 GTGAGACCGGAGCTGGAAGCATGACTTTGAGCCCTGATTCCTTTAAGGCTTCA 1087  
Db 841 GTGAAGAAAGAGGCCGAGAACATGCTTTTGGACCTTAGGTGAAGAAACATGCTTCA 899

Search completed: May 21, 2004, 20:33:04  
Job time : 5588 secs

Mon May 24 08:18:49 2004

us-09-931-457a-30.rng

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 21, 2004, 07:37:59 ; Search time 653 Seconds

(without alignments)  
8860.713 Million cell updates/sec

Title: US-09-931-457a-30

Perfect score: 1362

Sequence: 1 acttctgagtcgtagatag.....aaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 293a04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	1362	2	AAV99906
2	641.4	47.1	1120	3	AAC42360
3	636.4	46.7	1303	2	AAQ4450
4	618.2	45.4	946	3	ADCT6317
5	595.2	43.7	966	6	ABZ13088
6	595.2	43.7	966	7	ADA68423
7	545.2	40.0	993	7	ADA69412
8	529.4	38.9	5259	7	ADA70437
9	524.6	38.5	1483	2	AAQ74413
10	509.6	37.4	1503	3	AAC34662
11	497.4	36.5	1303	3	AAC33559
12	495.8	36.4	975	9	ADCG4221
13	495.8	36.4	1200	3	AAAC44727
14	486.4	35.7	1302	3	AAC45323
15	482	35.4	725	9	ADDI17591
16	479.2	35.2	1148	3	AAAC7965
17	478.8	35.2	1224	3	AAC40774
18	477.2	35.0	972	6	ABZ12690
19	477.2	35.0	999	3	AAAC5719
20	476.4	35.0	1329	6	ABZ12294
21	475.6	34.9	1423	3	AAAC1206
22	475.2	34.9	1215	3	AAAC4993
23	447.6	32.9	1030	9	ADDI17590

24	425	31.2	683	9	ADCT75695	ADCT5695 DNA homol
25	419	30.8	1037	9	ADCT6316	ADCT6316 DNA homol
26	416.2	30.6	918	3	AAC41635	AAC41635 Arabidops
27	415.6	30.5	1006	9	ADCT5697	ADCT5697 DNA homol
28	408.2	30.0	614	9	ADCT5699	ADCT5699 DNA homol
29	408.2	30.0	614	9	ADCT5699	ADCT5699 DNA homol
30	383.2	28.1	933	9	ADDI16769	ADDI16769 DNA (Seqi
31	381.6	28.0	933	9	ADDI16769	ADDI16769 DNA (Seqi
32	375.8	27.6	657	8	ADCT76310	ADCT76310 DNA homol
33	369	27.1	894	9	ADCT76316	ADCT76316 DNA homol
34	369	27.1	894	9	ADCT76316	ADCT76316 DNA homol
35	369	27.1	1431	3	AAC46837	AAC46837 Arabidops
36	360.2	26.4	1200	7	ADCT17130	ADCT17130 DNA clone
37	358.8	26.3	1200	7	ADCT17130	ADCT17130 DNA clone
38	357.5	26.3	1417	3	AAC51477	AAC51477 Rice gene
39	356.2	26.2	1481	3	AAC47004	AAC47004 Arabidops
40	356.2	26.2	1529	4	ABA01246	ABA01246 O-acetyl
41	337.4	24.8	1089	3	AAC51458	AAC51458 Arabidops
42	337.2	24.8	606	8	ADCT17123	ADCT17123 DNA clone
43	335.2	24.6	528	7	ABX57527	ABX57527 Arabidops
44	333	24.4	506	9	ADDI16772	ADDI16772 DNA (Seqi
45	330.4	24.3	675	8	ADCT17144	ADCT17144 DNA clone

#### ALIGNMENTS

RESULT 1	AAV99906	standard; cDNA; 1362 BP.
AAV99906	AAV99906	standard; cDNA; 1362 BP.
XX	AAV99906	
AC	AAV99906	
XX	AAV99906	
DT	27-SEP-1999	(first entry)
XX	27-SEP-1999	(first entry)
DE	Fragment of cysteine synthase gene.	
XX	Biosynthesis; biosynthetic pathway; lysine; threonine; methionine;	
KW	cysteine; isoleucine; amino acid; homoserine kinase;	
KW	aspartic semialdehyde dehydrogenase; diaminopimelate decarboxylase;	
KW	cysteine synthase; cystathione beta-lyase; gene expression; screening;	
KW	inhibition; ss.	
XX		
OS	Glycine max.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	90..1067
FT		/*tag= a
FT		/product= "Cysteine synthase fragment"
XX		
PN	W09856935-A2.	
XX	17-DEC-1998.	
PD	17-DEC-1998.	
XX		
PF	11-JUN-1998;	98WO-US012073.
XX		
PR	12-JUN-1997;	97US-0049406P.
XX		
PR	12-NOV-1997;	97US-0065385P.
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E. I.	
XX		
PI	Falco SC, Allen SM, Thorpe CJ;	
XX		
DR	WPI: 1999-080910/07.	
XX		
DR	P-PsDB: AAM81018.	
XX		
PT	New isolated plant amino acid biosynthetic enzyme nucleic acids - which	
XX	encode aspartic semialdehyde dehydrogenase, diaminopimelate	
PT	decarboxylase, homoserine kinase, cysteine synthase and cystathionine	
XX	beta-lyase.	
XX		
PS	Claim 17; Page 57-58; 80pp; English.	

CC Organisation of the pathway leading to plant biosynthesis of lysine,  
CC threonine, methionine, cysteine and isoleucine suggests that over-  
CC expression or reduction of expression of genes encoding enzymes involved  
CC in that biosynthetic pathway could be used to alter the level of these  
CC amino acids in human food and animal feed. This may increase the  
CC nutritional quality of human food and animal feed by increasing the  
CC production and accumulation of specific free amino acids. The enzymes  
CC include aspartate semialdehyde dehydrogenase, homoserine kinase,  
CC diaminopimelate decarboxylase, cysteine synthase and cystathione beta-  
CC lyase. The nucleic acids encoding these enzymes can be used for altering  
CC the level of expression of the enzymes and for evaluating compounds for  
CC their ability to inhibit the enzymes' activity  
XX

SQ Sequence 1362 BP; 388 A; 243 C; 330 G; 401 T; 0 U; 0 Other;

Query Match 100.0%; Score 1362; DB 2; Length 1362;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTGTAGTTCGATAGTACCCGATGCTTGTCTTATGTCGATGATTCCTGTTCT 60  
DB 1 ACTTGTAGTTCGATAGTACCCGATGCTTGTCTTATGTCGATGATTCCTGTTCT 60  
QY 61 CAAGTCAAGCTTTGATGAGCAAGATTAATGCTGTTGAAAGTCCGAATGCCAAG 120  
DB 61 CAAGTCAAGCTTTGATGAGCAAGATTAATGCTGTTGAAAGTCCGAATGCCAAG 120  
QY 121 AGGTATAGGAAATGATGTTGTTAAACCCATTAGTATCTTAATTAATCTGCGATGTT 180  
DB 121 AGGTATAGGAAATGATGTTGTTAAACCCATTAGTATCTTAATTAATCTGCGATGTT 180  
QY 181 GTGTGCGCGGGTGTCTGCTAACTGAGTTGATGAGCACTGCTCTGATGTTGAGGAC 240  
DB 181 GTGTGCGCGGGTGTCTGCTAACTGAGTTGATGAGCACTGCTCTGATGTTGAGGAC 240  
QY 241 GGATGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB 241 GGATGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 301 GTGTCTCATTTGAGCCCAAGTGTATATCTGCGATTTGATGATGATGATGATGATGAT 360  
DB 301 GTGTCTCATTTGAGCCCAAGTGTATATCTGCGATTTGATGATGATGATGATGATGAT 360  
QY 361 CCAGGGGTTCAGGCTGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 361 CCAGGGGTTCAGGCTGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 421 TTCTATAGCTTTTGGAGCTGAGTTGTTCTGACAGATCTGCTTAAGGAAATGAAAGT 480  
DB 421 TTCTATAGCTTTTGGAGCTGAGTTGTTCTGACAGATCTGCTTAAGGAAATGAAAGT 480  
QY 481 CTGTTCAGAGGCTGAGAGATATTTGGCTTAAGAGCCCAATGCTTACATCTTCACAAAT 540  
DB 481 CTGTTCAGAGGCTGAGAGATATTTGGCTTAAGAGCCCAATGCTTACATCTTCACAAAT 540  
QY 541 TTGAAAGCCCTGCAATCCCAAGGTTCAATATGAAACCACTGCTCCAGAGATATGAAAG 600  
DB 541 TTGAAAGCCCTGCAATCCCAAGGTTCAATATGAAACCACTGCTCCAGAGATATGAAAG 600  
QY 601 GCTCCGATGGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 601 GCTCCGATGGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 661 GTGCTGAGAAATATCTTAAGAGCAGATCCGATATTAAGTGTGTTGAGAACAG 720  
DB 661 GTGCTGAGAAATATCTTAAGAGCAGATCCGATATTAAGTGTGTTGAGAACAG 720  
QY 721 TTGAAAGTCCAGTCTCTCAAGAGGAAAGCTGCTCAACAGATCAAGGATTTGTT 780  
DB 721 TTGAAAGTCCAGTCTCTCAAGAGGAAAGCTGCTCAACAGATCAAGGATTTGTT 780  
QY 781 CTGCTTTATCCCTGGGTCTTGAAGCAATCTTCTTGAAGTGTCAATATCA 840

DB 781 CTGCTTTATCCCTGGGTCTTGAAGCAATCTTCTTGAAGTGTCAATATCA 840  
QY 841 GGATGAGAGCAATAGAAAGTCCGATGCTTGGCTTAAAGAGGCTATTTGGGAA 900  
DB 841 GGATGAGAGCAATAGAAAGTCCGATGCTTGGCTTAAAGAGGCTATTTGGGAA 900  
QY 901 TATCTCCGAGCTCAGCTGCTGCTTTTACATTTGACATTTGACAAAGAACCAAAATGCCG 960  
DB 901 TATCTCCGAGCTCAGCTGCTGCTTTTACATTTGACATTTGACAAAGAACCAAAATGCCG 960  
QY 961 GGAAGCTATTTGTCGCTTTTCCAGCTTCCGAGAGGATGATGATGATGATGATGAT 1020  
DB 961 GGAAGCTATTTGTCGCTTTTCCAGCTTCCGAGAGGATGATGATGATGATGATGAT 1020  
QY 1021 TTGAGTCAAGTGAAGAGCGGAGGCTGAAAGCATGATGATGATGATGATGATGAT 1080  
DB 1021 TTGAGTCAAGTGAAGAGCGGAGGCTGAAAGCATGATGATGATGATGATGATGAT 1080  
QY 1081 GCTTCACTACATGATTTTCTGTTACTGTTACAGGCTTTTACATGATGATGATGAT 1140  
DB 1081 GCTTCACTACATGATTTTCTGTTACTGTTACAGGCTTTTACATGATGATGATGAT 1140  
QY 1141 TACTGTTGTAAGTCACTGATCTTAATTAATTAATTAATTAATTAATTAATTAAT 1200  
DB 1141 TACTGTTGTAAGTCACTGATCTTAATTAATTAATTAATTAATTAATTAATTAAT 1200  
QY 1201 GCCCTGAGAGCAATATTTGCTGCTGCACTTTAAAGATTTCAAGTGTCTTATAG 1260  
DB 1201 GCCCTGAGAGCAATATTTGCTGCTGCACTTTAAAGATTTCAAGTGTCTTATAG 1260  
QY 1261 TAAAGATTTTCACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1320  
DB 1261 TAAAGATTTTCACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1320  
QY 1321 TTCTGATTTCTTTGCTGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 1362  
DB 1321 TTCTGATTTCTTTGCTGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 1362

RESULT 2  
ID AAC42360 standard; DNA; 1120 BP.  
AC AAC42360;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DB Arabidopsis thaliana DNA fragment SEQ ID NO: 35256.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX  
PN BP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
XX  
PR 05-MAR-1999; 99US-0123180P.  
XX  
PR 09-MAR-1999; 99US-0123548P.  
XX  
PR 23-MAR-1999; 99US-0125788P.  
XX  
PR 25-MAR-1999; 99US-0126264P.  
XX  
PR 29-MAR-1999; 99US-0126785P.  
XX  
PR 01-APR-1999; 99US-0127462P.  
XX  
PR 06-APR-1999; 99US-0128234P.  
XX  
PR 08-APR-1999; 99US-0128714P.  
XX  
PR 16-APR-1999; 99US-0129845P.  
XX  
PR 19-APR-1999; 99US-0130077P.  
XX  
PR 21-APR-1999; 99US-0130449P.

Mon May 24 08:18:49 2004

us-09-931-457a-30.rng

Page 3

PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155658P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158023P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.



PR 14-OCT-1999; 99US-015637P.  
PR 14-OCT-1999; 99US-015638P.  
PR 18-OCT-1999; 99US-015958AP.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match	47.1%;	Score 641.4;	DB 3;	Length 1120;
Best Local Similarity	79.1%;	Pred. No. 68-169;		
Matches 762;	Conservative 0;	Mismatches 201;	Indels 0;	Gaps 0.

Oy	105	TCGCGAATTGCCAAGATGTTACGGGAATTGATTGGTATAAACCCTTGTGTATCTAAAT	164
Dd	158	TCGAATAATTGCTAAAGATGTCATGATTAATTGGGACATCCCATTTGGTATTGTAAC	217
Oy	165	AAACTTGCAGATGTTGTGTGTCGCCGGTGTCTGTAAACTGAGTTGATGAGCATGC	224
Dd	218	AATGTGTCTGAAGGATGTGTGGTGTGTGTGTGTCTGCTAACTTGATGATGAAACCGTGC	277
Oy	225	TCTAATGTGAAGACAGATATGGGTATAGTATGATTTGCTGATGACGAAGAGAGGACTT	284
Dd	278	TCTACGCTCAAGACAGGATGTGGTTTACATGATTTCTGATGCGAAAAGAGGCTTT	337
Oy	285	ATCACACCTGGAAGAGTGTCTCTATTGAGCCCAACAGGTGTAATATATGCAATTGATTA	344
Dd	338	ATCAAAACGAGGAGAGTGTGCTGATTAAGCCCAACAGTGAAAACATGTGAGTTGCGTTA	397
Oy	345	GCCCTCAATGCGACGACCGGCGTTACAAGCTCATATTAACAATGCGCTTCTATAGT	404
Dd	398	GCATTCAAGCGACGCTCCAAAGCTTACAGCTTATATTAACAATCGACGTTCTATAGT	457
Oy	405	CTTGAGAGAGATCAATCTATTATTAAGCTTTTGGAGCTGATGTGTTCTGACAGATCTGCT	464
Dd	458	ACTGAGAGAGATCAATCTATTACTTTTGGAGTGTGATGTGGTTTAACTGACCCACT	517
Oy	465	AAGGAATGAAGAGTGTCTGTACAGAGGCTGAAGATATTTGGCTAAGACGCCCAATGC	524
Dd	518	AAGGCGATGAAGAGGCTATTCGCAAAAGCGAAGAGATTTTGGCATAAACCCCATGT	577
Oy	525	TACATATCTCAACAAATTTGAAAACCTGTCCCAATCCCAAGTTCAATTGAAACACTGCT	584
Dd	578	TACATGTCTACAGCAGCTTTGAGAACCTGTCCCAACCTTAAGATCACTATGAGCTACGGGA	637
Oy	585	CCAGGATATGGAAGAGGCTCCGATGCGGAAAAATGATGCAATTTGTTCTGGGATAAGCACT	644
Dd	638	CTTGAGATATGGAAGAGCCTGTGTGCAAAATCGATGCTTTGTTCGTGGATTTGTACT	697
Oy	645	GCTGTACATPACACGCTGTGGAATAATATCTTAAAGACAGATCCCAATATTAAGCTG	704
Dd	698	GCTGTACCATTAACGCTGTGGAAGATCTTAAAGACAGAAAGCAAACTCTCAAGCTG	757
Oy	705	AATGTGTGGAACGATTTGAAGTCCAGTCTCTCAGAGAGAAAGCTGTCTCACACAG	764
Dd	758	TATGAGTGAAGCCAGCTTGAAGTCTATCTATCGGTGGGAACCGAGTCTCTCACAAAG	817
Oy	765	ATTCAAGGATTTGCTGTGTTTATCCCTGTGTCTTGAAGTCAATCTTTGATGAA	824

Db	818	ATTCAAGGATAGAGCGTGGTTTATTCACAGTGATATGTAATGTGATCTTATTGACGAA	877
Qy	825	GTTTGCAAAATATCAAGTGATGAAGCATATAGAACTGCAAGCTTCTTGCGCTTAAAGAA	884
Db	878	GTTGTTCAGGTTTCAGATGATGAATCCATTGACATGGCAAGGCGACTTGCTTTAAAGAA	937
Qy	885	GACCTATTGTGGAAATATCTTCCGAGAGCTGCAGCTGCTGCTTTTCAGATTGCAAA	944
Db	938	GGCTCTTCTTGGGAAATATCATTCGGTCAGACACTGCTGAGGAATTTAACTTGCAAG	997
Qy	945	AGACAGAAATGCGCGGAGCTTATTTGTCGCGTTTTCACAGCTTCGGGAGAGGTAC	1004
Db	998	AGCGCAGAAAACCTGGGAGCTTATTTGTGCGATATTTCCGAGTTTCGGGAGAGGTAT	1057
Qy	1005	CTGTGCTCCGTCTATTGTGAATCAGTGAAGAGCGAAAGTGAAGAAGACTTTTGAGAGCC	1066
Db	1058	CTATTCAGCGTACTTTTCGATGCGACAAGAAAGAGGAGCACTGATGACTTGAGAGCT	1117
Qy	1065	TGA 1067	
Db	1118	TGA 1120	

RESULT 3  
AAQ44450  
ID AAQ44450 standard; cDNA to mRNA; 1303 BP.  
XX

AC	AA044450;
XX	
DT	12-OCT-1994 (first entry)
XX	
DE	Cysteine synthase.
XX	
KW	Cysteine synthase; plant; expression; probe; ss.
XX	
OS	Splnacia cleracea.
XX	
FE	Key
FT	Location/Qualifiers
FT	52..1029
FT	/*tag= a
PN	
PD	JP06038770-A.
XX	
PD	15-FEB-1994.
XX	
PE	05-FEB-1992; 92JP-00020315.
XX	
PR	05-FEB-1992; 92JP-00020315.
XX	
PA	(MITS ) MITSUBISHI CORP.
PA	(MITU ) MITSUBISHI KASEI CORP.
XX	
DR	MP1; 1994-094834/12.
DR	P-PSDB; AAR49830.
XX	
PT	Novel gene coding cysteine synthase - used to increase the cysteine content of an agricultural plant.
PS	Claim 2; Page 4-6; 6pp; Japanese.
XX	
CC	The cysteine content in an agricultural product can be increased by expressing the cysteine synthase in a plant. Probes V822 and V812 used in the isolation of the gene are given in AA044483-84
CC	
CC	
XX	
SQ	Sequence 1303 BP; 367 A; 242 C; 311 G; 383 T; 0 U; 0 Other;
QY	Query Match 46.7%; Score 636.4; DB 2; Length 1303;
	Best Local Similarity 77.7%; Pred. No. 1.8e-167;
DB	Matches 769; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
49	ATAATGCTGTGTAAGATCGAATGCGAAGATGTCGAAATGATGTAATAAC 146
	AAATGGTGGAGCAAGCCTTATTGCTTAAGATGTGACTGATGATTCGAAAAACG 108







DT		20-NOV-2003 (first entry)
XX		
DE	Rice gene, SEQ ID 2735.	
KM	Plant; bacterial infection; fungal infection; viral infection; rice;	
KX	gene; ds.	
XX		
OS	Oryza sativa.	
PM	MO200300898-A1.	
XX		
ED	03-JAN-2003.	
XX		
PF	22-JUN-2001; 2001WO-IB001105.	
XX		
PR	22-JUN-2001; 2001WO-IB001105.	
XX	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
PZ		
PI	Chang H, Chen W, Cooper B, Glaesbrook J, Goff SA, Hou Y;	
P1	Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;	
DR	WPI; 2003-175290/17.	
PT	Identifying at least one gene involved in plant resistance or response to	
PP	pathogenic infection for conferring resistance or tolerance to a plant	
PT	bacterial, fungal or viral infection by determining or detecting plant	
PF	gene expression.	
PS	Claim 6; SEQ ID NO 2735; 899pp; English.	
XX		
CC	The present invention relates to a method (M1) for identifying genes	
CC	involved in plant resistance or response to pathogenic infection. M1	
CC	comprises identifying a gene whose expression is significantly altered in	
CC	the incompatible interaction of plant gene expression relative to	
CC	expression of the gene in an uninfected plant, in a mutant plant that	
CC	does not express a gene associated with response to pathogenic infection,	
CC	or in a corresponding incompatible or compatible interaction. (M1) is	
CC	useful for conferring resistance to resistance or tolerance to a plant to	
CC	bacterial, fungal or viral infection. The present sequence was used to	
CC	illustrate the invention.	
SQ	Sequence 993 BP; 263 A; 207 C; 290 G; 233 T; 0 U; 0 Other:	
XX		
XX		
Query Match	40.0%; Score 545.2; DB 7; Length 993;	
Best Local Similarity	73.4%; Pred. No. 5.6e-142;	
Matches 697; Conservative	0; Mismatches 253; Indels 0; Gaps 0	
QY	111 ATGCGAAGATGTTCACGGAAATTGGTATAAACCCCATTTAGTATCTTAATAACTT 170	
DB	40 ATCGCCHAGAATGTCAACCGAGTTGATTTGGGAACAOCGCCGTBTGACTCAACGGGGTG 99	
QY	171 GCGGATGTTGTGTGCCCGGGTGTCTGCTAAACTGGAATTGATGAGGCATGCTTACT 230	
DB	100 ACGGATGGGTGTGCTCGGGCGCGTCCGGGCAAGCTCGAATCATGTGAGCCATGCTCAC 159	
QY	221 GGGAAGGACAGGATTTGGTATGATGATGCTGATGTCAGAGAGAGAGGACTTATCACA 290	
DB	160 GTCAAGGATTTGATTTGATACATGATACCTGATGACAGAGAGAGAGGGGCTGATCAT 219	
QY	291 CCTGGAAGAGTGTCTCATTTGAGCCACAGTGTGTAATCTGGCATTTGATTAAGCTTC 350	
DB	220 CCAGGCAAGAGTGTCTGATTTGAGCCAACTAGTGGCAACAGGGCAATTGACTGGCTTC 279	
QY	351 ATGGGAGGAGCCAGGGGTTACAAAGCTCATTAATTAACATGCTGCTCTTAGATCTTGAG 410	
DB	280 ATGTGCTGTCGCAAGGTTTACAGGCTTTGATCTGATCAAGATGCGGCTTCAGAGCATGGAG 339	
QY	411 AGAAGATCATTTCTAATTAAGCTTTTGAAGCTGAGTTGCTTGTGACAGATCTGCTAAGGA 470	
DB	340 AGAGAGATCAATATGTAAGGCTTTTGTGCTGGAATTAATTAATCAACCACCTTTGGGA 399	
QY	471 ATGAAAGAGCTGCTTACAGAGGCTGAAAGAGATTTGGCTTAAGAGCCCAATGCTTACATA 530	

Dd		400	ATGAAAGAGACTGTCCAAAAGCAGAGAAGTCTGCACGCAAGAACAAACAATTATTC	459
Oy		531	CTTCACAAATTTGAAAAACCCTGCCAATCCCAAGTTCAATTATGAACCATCTGGTCCAGAG	590
Dd		460	CTCCAAACAATTTCAGAAACCCTGCTAACCCCAAAGTCAATTACGAGACACTGGACCTTGAA	519
Oy		591	ATAAGAAAGGCTCCGATGGGAAAAATGATGCAATTTGTTCTGGGATAGGCACCTGCTGT	650
Dd		520	ATCTGAAGAGAAACGAGAGGTAAAGTTGAATGTTTAAGTTTCTGTGTAATTTGGGACAGTGCC	579
Oy		651	ACAATPAACAGGTGCTGGAAAAATATCTTPAAGACGAAATCCGAATPATPAAAGCTATTGGT	710
Dd		580	ACTATTAATCTGAGCTCGACGATACCTCAGAGAGCAAAATCCTGATATCAAGATCTATGGT	639
Oy		711	GTCGAACCTAGTTGAAAGTCCAGTGCCTCTCACAGAGAAAGCTGGTCCACACAAGATTCAA	770
Dd		640	GTCGAGCCAGTCGAGAGGCTGTCTTATCTGATGAAAGCCTGGGCCACACAGATTCGA	699
Oy		771	GGCATTTGGTGTCTGGTTTTATCCCTGAGTCTTGGAGTCAATCTTCTTGATGAAGTTGT	830
Dd		700	GGATTTGAGAGCTGGTTTGTTCCTGGGGTCTGTGATGTGACCTCATCAATGAAGAACTGA	759
Oy		831	CAATATCAAGTAGTAAGACAAATGAAACTGCAAAAGCTTCTTGCGCTTAAAGAGGCTCTA	890
Dd		760	CAAGTTTCAAGTAGTAAGACTATCAGATGGCAAAAGCTCTTGCAATTGAAAGAGGGTTG	819
Oy		891	TTTGTCGGAATATCTTCCGAGCTGCAGCTGCTGCTCTTTCAGATTGCAAAAGACA	950
Dd		820	CTGTTGGAATATCTTCAAGTGCAGCTGCAGACAGACTGTGAAGCTTCGTCAGAGGCTG	879
Oy		951	GAATAATCCGAGAAAGCTTATTTGTTGCCGTTTTTCCAGCTTCGAGAGAGGTACTCTGTC	1010
Dd		880	GAATAATGAAGAAAACTTTTGTGTGTCTTCCCAAGCTTGGTGAAGCGGTACTCTTTCG	939
Oy		1011	TCCGTGCTATTTTGAGTCAAGTGAAGACGCAAGACTTAAAGCATGACTTTTGA	1060
Dd		940	TGGTGCTCTTCCAGTCCATCAAGAGAGACTTAAACATGATGGTTGA	989
AD				
RESULT 8				
ADA70437				
ID	ADA70437 standard; DNA; 5255 BP.			
XX				
AC	ADA70437;			
XX				
DT	20-NOV-2003 (first entry)			
XX				
DE	Rice gene, SEQ ID 3760.			
XX				
KW	Plant; bacterial infection; fungal infection; viral infection; rice;			
gene; ds.				
XX				
OS	Oryza sativa.			
XX				
PN	MO2003000898-A1.			
PD	03-JAN-2003.			
XX				
PF	22-JUN-2001; 2001WO-IB001105.			
XX				
PR	22-JUN-2001; 2001WO-IB001105.			
XX				
(SYGN ) SYNGENTA PARTICIPATIONS AG.				
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,			
PI	Katagiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;			
XX				
DR	WPI; 2003-175296/17.			
PT	Identifying at least one gene involved in plant resistance or response to			
PT	bacterial infection for conferring resistance or tolerance to a plant to			
PT	bacterial, fungal or viral infection by determining or detecting plant			



Mon May 24 08:18:49 2004

us-09-931-457a-30.png

Page 10

Qy	6	AACTTGGTAGAGAGATATATATGGCTGTGAAAGGCCGGAATTGCCAAATGTGA	126
Dp	216	AAAGCTGTGTCCTTTGGCTCCATCCACATTTGAAGGCTTAACTTCTGAAATGTTT	275
Qy	127	CGGAATTGATTTGGTAAACCCTCATAGTATATCTAATTAACCTTGGAGTGGTTGTG	186
Dp	276	CTCAGCTAATTGGAAAAACCCCAATGGTGTATCTCAACAATGATCAAAAGATCAAGTTG	335
Qy	187	CCCGGGTTCGTCTAAACTGGAGTTGATGGACCAATGGCTAGTGTGAAGACAGATTG	246
Dp	336	CAACATTTGCTGGAAAGCTTGAGCATGGAACCTTGCGAGTTCAGAGACAGAAATTG	395
Qy	247	GGTATAGTATGATTTGGCTGATGACGAAGAGAGGGACTTATCAACTTGGAAAGTGTCC	306
Dp	396	CTTACAGTATGATTTGATGATGCTGACGACAAAGAGATTATCAACTTGAAGACTATCC	455
Qy	307	TCAATTGACCAACAAGTGTAAATCTGGCAATTGGAATTAGCTTCATGACACACCAAGG	366
Dp	456	TAGTGGAGCTTACGAGTGGGAAATACGGAATATGAACTTGGCTTTATTAAGCTCTGCGAG	515
Qy	367	GTTACAGCTCATATATACATGCTCCGTCTTATGATGCTTGAGAGAGAAATCATTTCTAT	426
Dp	516	GATCCAAAGATTACTTTCACATCCACAGCTTCATGATGATGGAAGAGAGATTATCTTGA	575
Qy	427	TAGCTTTTGAGAGCTGAGTTGGTTCTGACAGATCTGCTTAAGGAATGAAGTGTCTTC	486
Dp	576	AAGCATTTGAGAGCTGAGTTGGTCTGATCGATCCAGCTAAGGGAATGAAGACAGATG	635
Qy	487	AGAAGCTGAAAGATTTGGCTTAAGAGCCCAATGCCATCATCTTCAACATTTTGAA	546
Dp	636	AGAAGGCTGAAAGAAATTTGAAABAACTTCGATTCCTACATGCTTCAGCAGTTTGCA	695
Qy	547	ACCCTGCCAATCCCAAGSITCATTTATGAACAACATGCTCCAGATATGAAAGGCTCCG	606
Dp	696	ATCTCGAATTCCAAGATATCATTTACAGACAAACAGTCCGAGATCTGGGAAGACAA	755
Qy	607	ATGGAAAATTTGATGCAATTTGTTCTTGGGATAGGCACTGGTGTGACAPAAAGGTGTG	666
Dp	756	AAGCAAAAGTGACATTTTGTGTGACAGGCATTTGGAATGGAACATTCGTGAGTTG	815
Qy	667	GAATATCTTAAAGACAGATCCGAATATTAAGCTGATTTGGTGGAAACCACTTGA	726
Dp	816	GACGTAATCTCAAGAACGTAAACCTGATGTGAGGTAATTTGATATAGAACTACAGAA	875
Qy	727	GTCAGTGTCTCAGAGGAAAGCTGTGTCCACACAAATTTAAAGGATTTGGTGTGTT	786
Dp	876	GCAACATATCTTTCTGGTGGAAAGCTGTGTCCACAAAGTTTAAGGACTTGGAGTGT	935
Qy	787	TTATCCCTGGTCTTGAAGTCAATCTTGTATGATGATGATGTTCAATATCAATGATG	846
Dp	936	TTGTTCCAGCAATTTGATTTGGGTGTATGATGAATATTAAGGATATATATGAG	995
Qy	847	AAGCAATAGAACTGCAAGCTTCTTGGCTTAAAGAGGCTATTTGTGGGAATATCTT	906
Dp	996	AAAGCTAGAAATGGCAAAACCAATTTGGCAATGAAAGAGCTGTGTGGTGGCAATTTAT	1055
Qy	907	CCGAGCTGACAGCTGCTGCTCTTTCAGTTTCATTAAGAACACACAGAAATTCGCGGAAGC	966
Dp	1056	CTGAGGACAGCAGCTGTCTGAGCTCAGATTGTGTAAAGACCTGAAATATGAGAAAC	1115
Qy	967	TTATTTGTGCGGTTTTCCAGCTTGGGAGAGGTAAGCTGTCTTCGCTATTTAGAT	1026
Dp	1116	TTATTTGCTGTGTGTTCCAAAGCTTTGGTAGAGATACCTGTATCATTTGTGTCCAGT	1185
Qy	1027	CAGTGACAACGGAAGCTGAAGATGACTTTTGAAGCCCTGAATTCGCGTTAAGGCTCT	1086
Dp	1176	CTATTTGAGAAAGATGTGAATACTGAAGCCAGAAATGTGAAGATATTTCTTGAATAT	1235

PR 18-JUN-1999; 99US-0139463P  
 PR 18-JUN-1999; 99US-0139750P  
 PR 18-JUN-1999; 99US-0139763P  
 PR 21-JUN-1999; 99US-0139817P  
 PR 22-JUN-1999; 99US-0139899P  
 PR 23-JUN-1999; 99US-0140353P  
 PR 23-JUN-1999; 99US-0140354P  
 PR 24-JUN-1999; 99US-0140695P  
 PR 24-JUN-1999; 99US-0140823P  
 PR 29-JUN-1999; 99US-0140991P  
 PR 30-JUN-1999; 99US-0141287P  
 PR 01-JUL-1999; 99US-0141842P  
 PR 01-JUL-1999; 99US-0142154P  
 PR 02-JUL-1999; 99US-0142555P  
 PR 06-JUL-1999; 99US-0143390P  
 PR 08-JUL-1999; 99US-0143803P  
 PR 09-JUL-1999; 99US-0143920P  
 PR 12-JUL-1999; 99US-0143977P  
 PR 13-JUL-1999; 99US-0143977P  
 PR 14-JUL-1999; 99US-0143977P  
 PR 15-JUL-1999; 99US-0144005P  
 PR 16-JUL-1999; 99US-0144085P  
 PR 16-JUL-1999; 99US-0144086P  
 PR 19-JUL-1999; 99US-0144325P  
 PR 19-JUL-1999; 99US-0144331P  
 PR 19-JUL-1999; 99US-0144332P  
 PR 19-JUL-1999; 99US-0144333P  
 PR 19-JUL-1999; 99US-0144334P  
 PR 19-JUL-1999; 99US-0144335P  
 PR 20-JUL-1999; 99US-0144332P  
 PR 20-JUL-1999; 99US-0144332P  
 PR 20-JUL-1999; 99US-0144332P  
 PR 21-JUL-1999; 99US-0144814P  
 PR 21-JUL-1999; 99US-0145086P  
 PR 21-JUL-1999; 99US-0145088P  
 PR 22-JUL-1999; 99US-0145085P  
 PR 22-JUL-1999; 99US-0145087P  
 PR 22-JUL-1999; 99US-0145089P  
 PR 22-JUL-1999; 99US-0145192P  
 PR 23-JUL-1999; 99US-0145192P  
 PR 23-JUL-1999; 99US-0145218P  
 PR 23-JUL-1999; 99US-0145224P  
 PR 26-JUL-1999; 99US-0145276P  
 PR 27-JUL-1999; 99US-0145313P  
 PR 27-JUL-1999; 99US-0145918P  
 PR 27-JUL-1999; 99US-0145919P  
 PR 28-JUL-1999; 99US-0145919P  
 PR 02-AUG-1999; 99US-0146386P  
 PR 02-AUG-1999; 99US-0146388P  
 PR 02-AUG-1999; 99US-0146389P  
 PR 03-AUG-1999; 99US-0147038P  
 PR 04-AUG-1999; 99US-0147204P  
 PR 04-AUG-1999; 99US-0147302P  
 PR 05-AUG-1999; 99US-0147192P  
 PR 05-AUG-1999; 99US-0147260P  
 PR 06-AUG-1999; 99US-0147303P  
 PR 06-AUG-1999; 99US-0147416P  
 PR 09-AUG-1999; 99US-0147493P  
 PR 09-AUG-1999; 99US-0147935P  
 PR 10-AUG-1999; 99US-0148171P  
 PR 11-AUG-1999; 99US-0148319P  
 PR 12-AUG-1999; 99US-0148341P  
 PR 13-AUG-1999; 99US-0148655P  
 PR 13-AUG-1999; 99US-0148684P  
 PR 16-AUG-1999; 99US-0149368P  
 PR 17-AUG-1999; 99US-0149175P  
 PR 18-AUG-1999; 99US-0149426P  
 PR 20-AUG-1999; 99US-0149722P  
 PR 20-AUG-1999; 99US-0149723P  
 PR 20-AUG-1999; 99US-0149929P  
 PR 23-AUG-1999; 99US-0149930P  
 PR 23-AUG-1999; 99US-0149930P  
 PR 25-AUG-1999; 99US-0150566P

PR 26-AUG-1999; 99US-0150884P  
 PR 27-AUG-1999; 99US-0151065P  
 PR 27-AUG-1999; 99US-0151066P  
 PR 30-AUG-1999; 99US-0151080P  
 PR 30-AUG-1999; 99US-0151303P  
 PR 31-AUG-1999; 99US-0151438P  
 PR 01-SEP-1999; 99US-0151930P  
 PR 07-SEP-1999; 99US-0152363P  
 PR 10-SEP-1999; 99US-0153070P  
 PR 13-SEP-1999; 99US-0153758P  
 PR 15-SEP-1999; 99US-0154018P  
 PR 16-SEP-1999; 99US-0154039P  
 PR 20-SEP-1999; 99US-0154779P  
 PR 22-SEP-1999; 99US-0155139P  
 PR 23-SEP-1999; 99US-0155466P  
 PR 24-SEP-1999; 99US-0155659P  
 PR 28-SEP-1999; 99US-0156458P  
 PR 29-SEP-1999; 99US-0156586P  
 PR 04-OCT-1999; 99US-0157117P  
 PR 05-OCT-1999; 99US-0157753P  
 PR 06-OCT-1999; 99US-0157865P  
 PR 07-OCT-1999; 99US-0158029P  
 PR 08-OCT-1999; 99US-0158232P  
 PR 12-OCT-1999; 99US-0158368P  
 PR 13-OCT-1999; 99US-0159293P  
 PR 13-OCT-1999; 99US-0159294P  
 PR 13-OCT-1999; 99US-0159295P  
 PR 14-OCT-1999; 99US-0159329P  
 PR 14-OCT-1999; 99US-0159330P  
 PR 14-OCT-1999; 99US-0159331P  
 PR 14-OCT-1999; 99US-0159637P  
 PR 14-OCT-1999; 99US-0159638P  
 PR 18-OCT-1999; 99US-0159584P  
 PR 21-OCT-1999; 99US-0160741P  
 PR 21-OCT-1999; 99US-0160767P  
 PR 21-OCT-1999; 99US-0160768P  
 PR 21-OCT-1999; 99US-0160770P  
 PR 21-OCT-1999; 99US-0160814P  
 PR 21-OCT-1999; 99US-0160815P  
 PR 22-OCT-1999; 99US-0160980P  
 PR 22-OCT-1999; 99US-0160981P  
 PR 22-OCT-1999; 99US-0160981P  
 PR 22-OCT-1999; 99US-0160981P  
 PR 25-OCT-1999; 99US-0161404P  
 PR 25-OCT-1999; 99US-0161406P  
 PR 25-OCT-1999; 99US-0161358P  
 PR 26-OCT-1999; 99US-0161360P  
 PR 26-OCT-1999; 99US-0161361P  
 PR 28-OCT-1999; 99US-0161920P  
 PR 28-OCT-1999; 99US-0161920P  
 PR 28-OCT-1999; 99US-0161993P  
 PR 29-OCT-1999; 99US-0162142P

Query Match 37.4%; Score 509.6; DB 3; Length 1503;  
 Best Local Similarity 70.0%; Pred. No. 6.3e-132;  
 Matches 686; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 93 GCTGTTAAAGATCGGATTCGCAAGATGTTACGGATTCGTTAAACCCCATTA 152  
 DB 360 GCTGTTAAGGAGCTTAATATGCGATTAAGCCCGCTACGCTATTGGAAAACTCCGATG 419  
 QY 153 GTATATCTAATAAATCTGCGATGTTGTTGTCGCCCGGTTGCTCTAAACTGAGTTG 212  
 DB 420 GTGACTTGAACATGATGATGCAAGGCTGTGTTGCAAGTGTGCTTAACCTTGAATC 479  
 QY 213 ATGAGCGATGCTCTAGAGTGAAGAGCAAGATGGGTATATGATGATTCGATGACAGA 272  
 DB 480 ATGAGACATGTTGAGGTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 539  
 QY 273 GAGAGGAGCTTATACACCTGGAAGAGTGTCTCATGAGCCAAAGATGTTATCT 332  
 DB 540 GAGAAAGACTTATATACCTGGAAGAGTGTCTCATGAGCCAAAGATGTTATCT 599



Mon May 24 08:18:49 2004

us-09-931-457a-30.ing

Page 12

QY 333 GGCATTGATTAAGCTTCAATGAGCAGCCAGGGTTACAGCTCATTAATACATGCTT 392  
|||  
DB 600 GGGATTGGCTTGATTCATTCATGCTGCTTCAAAAGGCTATTAATCTTGAAGATGCTT 659  
|||  
QY 393 GCTTCTATGATCTTGAAGAGAAATCATCTTATATGCTTTTGAGCTGATGTTGCTCG 452  
|||  
DB 660 GGCCTCATGATGTTGAAAGGCGGCTTCTTTTGAAGGCAATTTGAGCTGACCTTGCTTA 719  
|||  
QY 453 ACAGATCTCTGCTAAGGAATGAAGAGTGTCTTCAAGAGCTGAAGATTTGCTTAAG 512  
|||  
DB 720 ACTGAACCTGCAAAAGATATGATGATGAGCAATTCAGAGCTGAGAAATCTTGAAGAAA 779  
|||  
QY 513 AGCCCAATGCTTCACTACTTCAACATTTGAAAACCTGCTCCATCCCAAGTTCTAT 572  
|||  
DB 780 ACTCCCAATTCCTACATGCTCCCAAGCTTGAACACCTGCTCCCAATCCCAAGTTCTAT 839  
|||  
QY 573 GAAGCACTGCTCAGAGATATGAAGAGCTCCGATGGGAAATTTGATGATTTTCTT 632  
|||  
DB 840 GAGACATCTGCTCTGATTTGGAGATGACAGAGGCAAAATCATATTTGCTTGG 899  
|||  
QY 633 GGGATAGCACTGCTGCTACATTAACAGTGTCTGAAATATCTTAAGACAGAACTCG 692  
|||  
DB 900 GGGATAGCACTGCTGCTACATTAACAGTGTCTGAAATATCTTAAGACAGAACTCG 959  
|||  
QY 693 AATATAGCTGATTTGCTGAGAACAGATTAAGTCCAGTCTCAGAGAGAAAGCTT 752  
|||  
DB 960 GAATTAAGCTTATTTGCTGAGAACAGATTAAGTCCAGTCTCAGAGAGAAAGCTT 1019  
|||  
QY 753 GGTCCACACAGATTCAGAGATTTGCTGCTTTTATCCCTGCTGCTTGAAGTCAAT 812  
|||  
DB 1020 GAGCTCACAAGATTAAGAGATTTGAGAGCTGATTTGATCCCAAGATTTGATCTGCT 1079  
|||  
QY 813 CTCTTGATGAAGTGTCTTCAATATCAAGATGAAGCAATGAAGAACTGCAAGCTTCT 872  
|||  
DB 1080 AATGTGATGAATATCAATAGATTTCCAGTGAAGAACTGATTAAGAACTGCAAGCTTCT 1139  
|||  
QY 873 GCGCTTAAGAAAGAGCTTATTTGAGAAATCTTCCGAGCTGACGCTGCTGCTTCT 932  
|||  
DB 1140 GCTCTCAGAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199  
|||  
QY 933 CAGATTGCAAAAAGACCAAGAAATGCGGAGAGCTTATTTGCTGCTTCTTCCAGCTTCT 992  
|||  
DB 1200 CAGATTGCAAAAAGACCAAGAAATGCGGAGAGCTTATTTGCTGCTTCTTCCAGCTTCT 1259  
|||  
QY 993 GGGAGAGGATACCTGCTGCTGCTGCTTATTTGAGTCAAGTGAAGAGCAAGCTGAAGCATG 1052  
|||  
DB 1260 GGGAGAGGATACCTGCTGCTGCTGCTTCTTCAATGATGAGAGAGAGGCGCAAAATG 1319  
|||  
QY 1053 ACTTTGAGCCCTGAATTC 1072  
|||  
DB 1320 CAGCCGAGCTTGAATTTCT 1339  
|||

PE 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126284P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130442P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 14-MAY-1999; 99US-0134768P.  
PR 18-MAY-1999; 99US-0134941P.  
PR 19-MAY-1999; 99US-0135124P.  
PR 20-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 24-MAY-1999; 99US-0136021P.  
PR 25-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 07-JUN-1999; 99US-0138094P.  
PR 08-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139493P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.



QY 570 TATGAACCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 629  
 DB 573 TATGAACCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 632  
 QY 630 TCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 689  
 DB 633 GCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 692  
 QY 690 CCGAATATGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 749  
 DB 693 AAGATCTTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 752  
 QY 750 CCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 809  
 DB 753 CCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 812  
 QY 810 AATCTTCTGATGAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 869  
 DB 813 AAGATCTTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 872  
 QY 870 CCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 929  
 DB 873 CCGGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 932  
 QY 930 TTTGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 989  
 DB 933 AATGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 992  
 QY 990 TTTGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 1049  
 DB 993 GATGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 1052  
 QY 1050 ATGATCTTTGA 1060  
 DB 1053 TTGCAATTC 1063

RESULT 12  
 AD64221  
 ID AD64221 standard; cDNA; 975 BP.  
 AC AD64221;  
 DT 18-DEC-2003 (first entry)  
 DE Thalecress cysteine synthase coding region.  
 KM Thalecress; cysteine synthase; ss: O-acetyl-L-serine: sulphide;  
 KM coumarin dye; cysteine-coumarin conjugate; herbicide; fungicide;  
 KM bactericide; insecticide; plant.  
 OS Arabidopsis thaliana.  
 XX  
 XX Key Location/Qualifiers  
 FT 1 975  
 FT /\*tag= a  
 FT /product= "Cysteine synthase"  
 FT /EC\_number= "EC4.2.99.8"  
 XX  
 XX US2003087327-A1.  
 PD 08-MAY-2003.  
 PF 13-JUL-2001; 2001US-00905290.  
 PR 13-JUL-2001; 2001US-00905290.  
 PA (RICE/) RICE J.  
 PA (LANN/) LANNING B.  
 PA (CRANFORD J. CRANFORD J.  
 PA (NTEG/) NTE G.  
 PI Rice J, Lanning B, Crawford J, Nye G;

XX MPI, 2003-777218/73.  
 DR P-PSDB; AD64218.  
 XX  
 PT Determining cysteine synthase activity by contacting cysteine produced by  
 PT the enzyme in a reaction mixture with a coumarin dye, exciting the formed  
 PT conjugate and detecting the fluorescent light emitted by conjugate.  
 XX  
 PS Disclosure; SEQ ID NO 3; 32pp; English.  
 CC The invention relates to determining (M) cysteine synthase activity, by  
 CC combining O-acetyl-L-serine, sulphide and cysteine synthase to form  
 CC reaction mixture under conditions suitable for production of cysteine,  
 CC contacting reaction mixture with coumarin dye capable of forming cysteine  
 CC -coumarin conjugate, and subjecting the reaction mixture to ultraviolet  
 CC (UV) light to excite the conjugate, and detecting fluorescent light  
 CC emitted by the conjugate. The cysteine synthase is a plant cysteine  
 CC synthase appearing as AD64218. The method is useful for determining  
 CC cysteine synthase activity by quantitating the cysteine formed by the  
 CC enzyme and for identifying a test compound as a herbicide candidate  
 CC (which involves performing the method in the presence and absence of the  
 CC compound, and comparing the amount of the fluorescent light emission in  
 CC the presence and absence of the test compound, where a decrease in the  
 CC amount of the fluorescent light emission in the presence of the test  
 CC compound indicates that the test compound is a herbicide candidate. The  
 CC method is also useful for identifying compounds capable of selectively  
 CC inhibiting plant, fungal and/or bacterial cysteine synthase activity  
 CC (which involves performing the method in the presence of plant cysteine  
 CC synthase and determining the inhibitory activity of the compound based on  
 CC fluorescent light emission, repeating the steps using a fungal or  
 CC bacterial cysteine synthase, and identifying compounds that selectively  
 CC inhibit plant, fungal or bacterial cysteine synthase). The method is  
 CC useful for identifying a compound as a candidate for herbicides,  
 CC fungicides, bactericides and insecticides. The present sequence is the A.  
 CC thaliana cysteine synthase coding region.  
 XX  
 SQ Sequence 975 BP; 297 A; 187 C; 253 G; 238 T; 0 U; 0 Other;

Query Match 36.4%; Score 495.8; DB 9; Length 975;  
 Best Local Similarity 69.4%; Pred. No. 3,8e-128;  
 Matches 674; Conservative 0; Mismatches 297; Indels 0; Gaps 0;

QY 90 AATGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 149  
 DB 1 AATGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 160  
 QY 150 TTAGTATATCTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 209  
 DB 61 AATGATATCTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 120  
 QY 210 TTAGTATATCTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 269  
 DB 121 AATGATATCTGAAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 180  
 QY 270 GAAAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 329  
 DB 181 GAAAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 240  
 QY 330 ACTGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 389  
 DB 241 ACCGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 300  
 QY 390 CCGTCTTATGAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 449  
 DB 301 CCGTCTTATGAGTGTGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 360  
 QY 450 CTGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 509  
 DB 361 CTGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 420  
 QY 510 AAGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 569  
 DB 421 AAGAGATGAGCACTGGTCAGAGATATGAAAGCTCCGAGGGAATAATGATTTGTT 480

QY 570 TATGAACCTCTGTCAGAGATATGAAAGGCTCCGATGGAAATATGATTCATTGTT 629  
 DB 481 TATCAACCCGCGGTCACAGATATGAGAGACTCTGACGGGAAAGTAGACATATTTGTT 540  
 QY 630 TTGGGATAGGACATGCTGATCAATTAACAGCTGCGAAATATCTTAAAGACAGAT 689  
 DB 541 GCCGCTGTGGAACCTGCTGGAACGCTTCGAGATCAGAGATTTCTTCAAGAGAGAT 600  
 QY 690 CCGAATATTAAGCTGATGCTGCTGAGACCAATTAAGTCACTGCTCTCAGAGAGAG 749  
 DB 601 AAAGACTTAAAGTTATGATGCTGAGAACTTACAGAAAGTCCGCTATATAGTGAAGGCAA 660  
 QY 750 CCTGCTCCACAGATTTGAGGAGATGCTGCTGCTTATCCCTGCTGCTTGAAGTC 809  
 DB 661 CCGGCTACCATTTGATTCAGATGATGAGGCTGAGCTCACTCCCAACATTTGATTTTC 720  
 QY 810 AATCTCTGATGATGATGCTTCAATATCAAGTATGAGCAATAGAAACTGCAAGCTT 869  
 DB 721 AAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 QY 870 CTTCGCTTAAAGAGGCTTATTTGAGGATATCTTCGAGCTGAGCTGCTGCTGCT 929  
 DB 781 CTTCGCTTAAAGAGGCTTATTTGAGGATATCTTCGAGCTGAGCTGCTGCTGCTGCT 840  
 QY 930 TTTCAGATGCAAAAAGACCAAGAAAGCCGAGAGCTTATTTGCTTTTCCAGC 989  
 DB 841 ATTAAGTGGCAAGAGGCGCCAGAAAGCGCGGCAACTCATGTTGATTTTCCCTAC 900  
 QY 990 TTTCGAGAGAGTACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049  
 DB 901 GGTGGGAGAGCTTACCTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 1050 ATGACTTTGA 1060  
 DB 961 TTGCCATTCA 971

RESULT 13  
 AAC4727  
 ID AAC4727 standard; DNA; 1200 BP.  
 AC AAC4727;  
 AC AAC4727;  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 43907.  
 XX  
 KM Hybridisation assay; genetic mapping; gene expression control;  
 KM protein identification; signal transduction pathway; metabolic pathway;  
 KM promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128714P.  
 PR 08-APR-1999; 99US-0129845P.  
 PR 16-APR-1999; 99US-0130077P.  
 PR 19-APR-1999; 99US-0130449P.  
 PR 21-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140931P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 28-JUL-1999; 99US-0145918P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148319P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 25-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 13-SEP-1999; 99US-0154018P.  
 PR 15-SEP-1999; 99US-0154039P.  
 PR 16-SEP-1999; 99US-0154779P.  
 PR 20-SEP-1999; 99US-0155139P.  
 PR 22-SEP-1999; 99US-0155486P.  
 PR 23-SEP-1999; 99US-0155659P.  
 PR 24-SEP-1999; 99US-0156458P.  
 PR 28-SEP-1999; 99US-0156589P.  
 PR 29-SEP-1999; 99US-0156589P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.

PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 25-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 26-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 36.4%; Score 495.8; DB 3; Length 1200;  
 Best Local Similarity 69.4%; Pred. No. 4, 2e-128;  
 Matches 674; Conservative 0; Mismatches 297; Indels 0; Gaps 0;

90 ATGGCTGTGAAAGTCCGGAATTCGCCAAGATGTTACGAAATTAATGTTAAACCCCA 149  
 226 ATGGAGAGACCGGTGTGATCAAGATGATCAATTAATGTTAAACCCCA 285  
 150 TTAGTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 209  
 286 ATGCTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 345  
 210 TTGATGAGCCTGCTCTGATGTTGAAGGACAGATTTGGTATGATGATGATGATGATGAT 269  
 346 ATGATGAGCCTGCTCTGATGTTGAAGGACAGATTTGGTATGATGATGATGATGATGAT 405  
 270 GAAGAGAGACCTTATCACTGGAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 329  
 406 GAAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 465  
 330 ACTGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 389  
 466 ACCGAGATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 525  
 390 CCTGCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 449  
 526 CCTGCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 585  
 450 CTGACAGATCTGCTTAAGGAAATGAAGGCTGTTCAAGAGCTGAAGATTAATTAATTAAT 509  
 586 CTGACAGATCTGCTTAAGGAAATGAAGGCTGTTCAAGAGCTGAAGATTAATTAATTAAT 645  
 510 AAGACCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 569  
 646 AAGATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 705  
 570 TATGAACCACTGCTCAAGATTAATGAAGGCTGTTCAAGAGCTGAAGATTAATTAATTAAT 629  
 706 TATGAACCACTGCTCAAGATTAATGAAGGCTGTTCAAGAGCTGAAGATTAATTAATTAAT 765  
 630 TCTGAGATGAGCTGCTGTTCAATTAATGAAGGCTGTTCAAGAGCTGAAGATTAATTAAT 689  
 766 GCGGCTGTTGAATCTGCTGAAGGCTGTTCAAGAGCTGTTCAAGAGCTGAAGATTAATTAAT 825  
 690 CCGATATTAAGCTGATTTGTTGAAGGCTGTTCAAGAGCTGTTCAAGAGCTGAAGATTAAT 749  
 826 AAGATCTTAAGCTTATGTTGAAGGCTGTTCAAGAGCTGTTCAAGAGCTGAAGATTAAT 885  
 750 CTTGCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 809  
 886 CCGGCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 945

```
QY      810 AATCTTCTTGATGAGTGTCTCAATATCAAGTGAATGAGCAATAGAACTGCAAGCTT 869
DB      946 AACGTTCTTGATGAGTATCATCCAGTGAACAGTGGAGCAATGAAAGCCAACTT 1005
QY      870 CTGGGCTTAAAGAGGCTTATTTGTGGAAATCTTCGAGACTGAGCTGCTGCT 929
DB      1006 CTGGCCCGAAGAGAGATTAATCTGCTGGAAATATCTTCGAGCTGCTGAGCCGCTGCG 1065
QY      930 TTTGAGATGCAAAAGACCAAGAAATCCCGGAGAGCTTATTTGCTCCGTTTCCGACG 989
DB      1066 ATPAAAGTGGCAAAACCGCCGCAAAACCGCGCAACTCATAGTTGATTTCCCTAGC 1125
QY      990 TTGGGGAGAGGTACTGCTGCTCCGCTGCTATTGAGTCAAGTGAAGCCGAAAGCTGAAAGC 1049
DB      1126 GGTGGGGAACGTTACTCTATCACTCCTCATTGTCGATCAAGTCAAGATGAAGAGAGAAAT 1165
QY      1050 ATGACTTTTGA 1060
DB      1186 TTGCCAATTC 1196

RESULT 14
AAC45323
ID      AAC45323 standard; DNA; 1302 BP.
XX      AAC45323;
XX      18-OCT-2000 (first entry)
DB      Arabidopsis thaliana DNA fragment SEQ ID NO: 46095.
XX      Hybridisation assay; genetic mapping; gene expression control;
KM      protein identification; signal transduction pathway; metabolic pathway;
KW      promoter; termination sequence; ss.
OS      Arabidopsis thaliana.
XX      EP1033405-A2.
XX      06-SEP-2000.
XX      25-FEB-2000; 2000EP-00301439.
XX      25-FEB-1999; 99US-0121825P.
XX      05-MAR-1999; 99US-0123180P.
XX      09-MAR-1999; 99US-0123548P.
XX      23-MAR-1999; 99US-0125788P.
XX      25-MAR-1999; 99US-0126264P.
XX      29-MAR-1999; 99US-0126785P.
XX      01-APR-1999; 99US-0127462P.
XX      06-APR-1999; 99US-0128234P.
XX      08-APR-1999; 99US-0128714P.
XX      16-APR-1999; 99US-0129845P.
XX      19-APR-1999; 99US-0130077P.
XX      21-APR-1999; 99US-0130449P.
XX      23-APR-1999; 99US-0130510P.
XX      23-APR-1999; 99US-0130891P.
XX      28-APR-1999; 99US-0131449P.
XX      30-APR-1999; 99US-0132048P.
XX      04-MAY-1999; 99US-0132407P.
XX      05-MAY-1999; 99US-0132484P.
XX      06-MAY-1999; 99US-0132485P.
XX      06-MAY-1999; 99US-0132486P.
XX      06-MAY-1999; 99US-0132487P.
XX      07-MAY-1999; 99US-0132863P.
XX      11-MAY-1999; 99US-0134256P.
XX      14-MAY-1999; 99US-0134218P.
XX      14-MAY-1999; 99US-0134219P.
XX      14-MAY-1999; 99US-0134221P.
XX      14-MAY-1999; 99US-0134370P.
XX      18-MAY-1999; 99US-0134768P.
XX      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136382P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139859P.
PR      23-JUN-1999; 99US-0140354P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140892P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142399P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144088P.
PR      16-JUL-1999; 99US-0144088P.
PR      16-JUL-1999; 99US-0144088P.
PR      16-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144335P.
PR      21-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145086P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145089P.
PR      23-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145224P.
PR      23-JUL-1999; 99US-0145276P.
PR      26-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145918P.
PR      28-JUL-1999; 99US-0145951P.
```

Mon May 24 08:18:49 2004

us-09-931-457a-30.rng

Page 18

PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147433P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 22-SEP-1999; 99US-0155486P.  
PR 22-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159234P.  
PR 13-OCT-1999; 99US-0159234P.  
PR 13-OCT-1999; 99US-0159235P.  
PR 14-OCT-1999; 99US-0159332P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161935P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 35.7%; Score 486.4; DB 3; Length 1302;  
Best Local Similarity 69.4%; Pred. No. 1.9e-125;  
Matches 675; Conservative 0; Mismatches 296; Indels 1; Gaps 1;

90 ATGCTGTTGAAGGTCGGAATTCACAAAGATGTACGGAATGATGTGTTAAACCCCA 149  
91 ATGGAGGAGACCGGTGTGATCAAGATGATGACACTCAATTTGATGATTAACCCCA 150  
150 TTAGTATATCTAAATTAATCTGCGATGTGTGTGTCGCGGTGTGCTAACTGAG 209  
151 ATGTATATCTGAACACATTTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 210  
210 TTGATGAGCCATCTCTAGTGTGAAGA-CAGATTGGGTATGATGATTTGCTGATGC 268  
211 ATGATGAGCCCTGCTCTAGTGTCAAGAGGCGGATTCCTATGTATGATTAAGATGC 270  
269 AAGAGAGAGAGACTTATCAACCTGGAAGAGTGTCTCATTTGAGCAACAGTGTATA 328  
271 AAGAGAGAGAGATTTATCTCTGGAAGAGACACTGATTTGAGGCTTCTCTGTATA 330  
329 TACTGCAATGATGATGATTCATGCGACAGCCAGGCTTCAAGCTCATATTAAT 388  
331 CACCGGATGTGTTAGCTTCAATGCTGAGCTAAAGGTTACAAAGTGTCTCAAT 390  
388 GCTGCTCTTATGATCTTGTGAGAGAGATCATCTTATGATTTGAGCTGATGAT 448  
391 GCTCTATCAATGAGCTTGAAGAAATCTTTTATGATTTAGTGTGAGGATTTCA 450  
449 TCTGACATCTCTGTAAGGAGATGAAGATGCTGTCAGAAAGCTGAGATATTTG 508  
451 CCTCAAGATCTTATTAAGGCTTCAAGAAATATTCACAAAGCTGAAGATATTTAG 510  
509 TAAAGAGCCCAATGCTTATCTTCAACATTTTAAACCTGCAATCCCAAGTTCA 568  
511 CAATAATCAATGATATCAATGCTTCAACATTTTAAACCTGCAATCCCAAGTTCA 570  
569 TATGAAACCACTGCTCAGATATGAAGAGCTCCAGTGAAGAAATGATGATTTGT 628  
571 TATGAAACCAAGGCTTCAAGATATGAAGAGCTCCAGTGAAGAAATGATGATTTGT 630  
629 TTTGGATGAGCACTGCTGTATCAATTAAGAGCTGTTAAATTTTAAAGACAGA 688  
631 TCCGCTGTGTAAGCTGTTGAGAAAGCTTCCGATCAGAAATTTCTCAAGAGAGAA 690  
689 TCGAATATTAAGCTGATGTTGTTGAGAAACCAATGTAAGTCTCTCAGAGAGAA 748  
691 TAAAGCTTAAAGTTATGTTGTGAGAACTTCAAGAAAGTGTGATTAAGAGAGCA 750  
749 GCTGCTCACAAGATTTCAAGATTTGTTGTTTATTTTATTTTATTTTATTTTGA 808  
751 ACCGGGTACATTTGATTCAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 810  
809 CATCTCTTGAAGAGTTGTTCAATATCAATGATGAGAAATGAACTGCAAGCT 868  
811 CAGCTTCTGAAGATGATCAATGATGAGAAATGAACTGCAAGCTGCAAGCT 870  
869 TCTTGGCTTAAAGAGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 928  
871 TCTTGGCTTAAAGAGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 930  
929 TTTTGAATGAGAAAGAGCTTAAATGAGAAAGCTTATTTGTTGTTGTTGTTGTTG 988  
931 GATTAAGTGTGAGAAAGAGCTTAAATGAGAAAGCTTATTTGTTGTTGTTGTTGTTG 990  
989 CTTGGGAGAGTACTTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1048  
991 CGGTGGGAGAGCTTACTTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1050





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 11:41:31 ; Search time 3959 Seconds  
(without alignments)  
10273.375 Million cell updates/sec

Title: US-09-931-457A-30

Perfect score: 1362

Sequence: 1 acctgtgacgtcgtgacgtacg.....aaaaaaaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues 55026578

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_estbm:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rdx:\*  
26: em\_gss\_pkg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Query length	DB ID	Description
1	768.8	56.4	772	14	CAB01406
2	758	55.7	771	12	B1970016
3	754.4	55.4	776	12	B1969937
4	692.4	50.8	719	10	B858125
					B858125 GM700004B

# SUMMARY

5	651.4	47.8	803	12	BG648907	BG648907	EST510526
6	642	47.1	678	10	AM311624	AM311624	EST4307.Y
7	636.4	46.7	929	14	CF513886	CF513886	CABud0007
8	630.4	46.3	632	14	CAB01673	CAB01673	sat16d04.
9	629.2	46.2	910	14	CB974516	CB974516	CAB30005
10	627.2	46.0	724	10	BF003458	BF003458	EST431956
11	626.2	46.0	791	12	BG555822	BG555822	EST494500
12	625.4	45.9	994	14	CK260011	CK260011	EST706089
13	625	45.9	653	10	BH660317	BH660317	2-D7 Gmax
14	612.8	45.0	967	14	CK271210	CK271210	EST717288
15	606	44.5	613	13	BQ630428	BQ630428	sag07c05.
16	599	44.0	937	14	CK26154	CK26154	EST712312
17	595.6	43.7	785	12	BG587892	BG587892	EST488667
18	595.4	43.7	734	12	BG648870	BG648870	EST510489
19	588.2	43.2	673	12	BM815595	BM815595	EST593689
20	584.6	42.9	920	14	CK259857	CK259857	EST705935
21	581.6	42.7	907	14	CK260012	CK260012	EST706090
22	581.6	42.7	917	14	CK259858	CK259858	EST705936
23	578.4	42.5	934	29	CG848066	CG848066	ZMMBB031
24	577.2	42.4	935	14	CK279506	CK279506	EST725584
25	576.2	42.3	881	14	CK259489	CK259489	EST705567
26	573.4	42.3	655	12	B1309270	B1309270	NP100C02D
27	572.2	42.0	653	12	BG451772	BG451772	NP100C02D
28	571.8	42.0	873	14	CK259488	CK259488	EST705566
29	570.6	41.9	845	14	CF200608	CF200608	RR890915N
30	570.4	41.9	658	12	B1264457	B1264457	NP13301P
31	568.6	41.7	802	14	CB974696	CB974696	CAB30005
32	567.4	41.7	570	12	BM891119	BM891119	sam23b05.
33	566	41.6	745	12	BM815488	BM815488	EST593582
34	566	41.6	886	14	CK256667	CK256667	EST711745
35	562.2	41.3	792	14	CF512887	CF512887	CABud0005
36	562.2	41.2	813	14	CF205785	CF205785	RR890915I
37	559.2	41.1	686	12	B1264323	B1264323	NP118604P
38	557.4	40.9	1483	11	AY103651	AY103651	Zea mays
39	550.8	40.4	554	12	BM891219	BM891219	sam24d01.
40	546.6	40.1	709	14	CA411511	CA411511	B813.F-P
41	544.8	40.0	581	10	AM830316	AM830316	sm25d07.Y
42	544.6	40.0	623	10	AM775920	AM775920	EST734985
43	543.8	39.9	627	10	BP005244	BP005244	EST433742
44	543.4	39.9	856	14	CF652605	CF652605	65-102057
45	533	39.1	666	12	B1265217	B1265217	NP103E11I

## ALIGNMENTS

RESULT 1	CAB01406	772 bp	linear	EST 05-DEC-2002
LOCUS	sa05c03.y2	Gm-cl062	glycine max	CDNA clone SOYBEAN
DEFINITION	Gm-cl062-7397.5'	similar to	SW:CYSK_CITLA Q4317	CYSTEINE SYNTHASE
ACCESSION	CAB01406	GI:26058492		
VERSION	CAB01406.1			
KEYWORDS	glycine max (soybean)			
SOURCE	Glycine max			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.			
REFERENCE	1 (bases 1 to 772)			
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Espeland, J., Corry, V., Khanna, A., Bolla, B., Marita, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Steptoe, M., Thelting, B., Allen, M., Bowers, J., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, B., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.			
TITLE	Public Soybean EST Project			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine			

Mon May 24 08:18:50 2004

us-09-931-457a-30.rst

Page 2

4444 ForestPark Parkway, Box 9501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: [cuw@resgen.com](mailto:cuw@resgen.com) web site:  
[www.resgen.com](http://www.resgen.com)  
Seq. primer: -40RP from Glibco  
High quality sequence, step: 436.

**FEATURES**  
**source**

Location/Qualifiers  
1..772

/ncoes=Iwcoes; pBluescriptII SK<sup>+</sup>; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from stem tissue of 1 month old greenhouse grown plants for the cultivar Raden. Complementary DNA was synthesized from mRNA using a primer consisting of a synthesized (AT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10S host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

**ORIGIN**

Query Match 56.4%; Score 768.8; DB 14; Length 772;

Best Local Similarity 99.7%; Pred. No. 2.2e-130;  
Matches 770; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	59	CTCAAGTCAAGCCTTGTAGTAGGAGAGATATATGCTGTGTAAAGATCCGGAATTGGCAA	118
Db	1	CTCAAGTCAAGCCTTGTAGTAGGAGAGATATATGCTGTGTAAAGATCCGGAATTGGCAA	60
QY	119	AGATGTTACGGAAATTGATTTGGTAAACCCTATAGTATATCTAAATATTAATCTTGGGATGG	178
Db	61	AGATGTTACGGAAATTGATTTGGTAAACCCTATAGTATATCTAAATATTAATCTTGGGATGG	120
QY	179	TTTGTTGCCCGGGTTGCTGTAACTGGATGTATATGAGCAATGCTCTAGTGTGAAGA	238
Db	121	TTTGTTGCCCGGGTTGCTGTAACTGGATGTATATGAGCAATGCTCTAGTGTGAAGA	180
QY	239	CAGGATTTGGTATAGTATGATTGCTGTAGCAGAGAGAGAGGACTTATACACCTCGAAA	298
Db	181	CAGGATTTGGTATAGTATGATTGCTGTAGCAGAGAGAGAGGACTTATACACCTCGAAA	240
QY	299	GAGTGTCTCATTTGAGCCCAAAATGTGTAAATATCTGGCATTTGGATTAGCTTCATGGCAGC	358
Db	241	GAGTGTCTCATTTGAGCCCAAAATGTGTAAATATCTGGCATTTGGATTAGCTTCATGGCAGC	300
QY	359	AGCCAGGAGGTATACAAGCTCATATATTACAAATGCTGCTCTATAGTCTTTGAGAGAGAT	418
Db	301	AGCCAGGAGGTATACAAGCTCATATATTACAAATGCTGCTCTATAGTCTTTGAGAGAGAT	360
QY	419	CATTCTATAGCTTTTGGAGCTGAGTGTGTTCTGACAGATCTCTGTAAGGGAATGAAAGG	478
Db	361	CATTCTATAGCTTTTGGAGCTGAGTGTGTTCTGACAGATCTCTGTAAGGGAATGAAAGG	420
QY	479	TGCTGTTCAGAGAGCTGAAGAGATTTGGCTTAAGAGCGCCATGCTCTACATCTTCAACA	538
Db	421	TGCTGTTCAGAGAGCTGAAGAGATTTGGCTTAAGAGCGCCATGCTCTACATCTTCAACA	480
QY	539	ATTTGAATAACCTGCGCAATCCCAAGTTCATTATGAAACCATGTGCTCAGAGATATGGA	598
Db	481	ATTTGAATAACCTGCGCAATCCCAAGTTCATTATGAAACCATGTGCTCAGAGATATGGA	540

QY	599	AGGCTCCCATGGGAAAATGATGCAATTGGTTCTGCGGATAGGCACCTGATGGTCAATAAC	658
Db	541	AGGCTCCATGGGAAAATGATGCAATTGGTTCTGCGATAGGCACCTGATGGTCAATAAC	600
QY	659	AGGTGCTGGAAATATCTTAAAGACAGAAATCCGATATAAAGCTGATTGGTGTGAAC	718
Db	601	AGGTGCTGGAAATATCTTAAAGACAGAAATCCGATATAAAGCTGATTGGTGTGAAC	660
QY	719	AGTTGAAATCCAGTCTCTCAGAGGAAACCTGCTCCACAGATTCAGGGATTTGG	778
Db	661	AGTTGAAATCCAGTCTCTCAGAGGAAACCTGCTCCACAGATTCAGGGATTTGG	720
QY	779	TGCTGATTTTATCCCGGAGCTTGGAGTCAATCTTCTGATGAGGTGT	830
Db	721	TGCTGATTTTATCCCGGAGCTTGGAGTCAATCTTCTGATGAGGTAT	772

## RESULT 2

LOCUS	771 bp	mRNA	linear	EST 23-OCT-2001
DEFINITION	B1970016	Gm-r1083	Glycine max	CDNA clone Gm-r1083-3109 3'
DEFINITION	GM830009B12A07			
DEFINITION	mRNA sequence.			

**DEFINITION**

GM830009B12A07 Gm-r1083 Glycine max cDNA clone Gm-r1083-3109 3', mRNA sequence.

**ACCESSION**

VERSION

**KEYWORDS**  
coping

**SOURCE**  
**OPICANTSM**

## ORGANISM

## REFERENCE

## AUTHORS

**TITLE**

## JOURNAL

**COMMENT**

1 (bases 1 to 773)  
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V.,  
Expanding, J., Ralph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Other ESTs: AW30316 corresponding to Gm-Cl028-4934 (5')  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu  
This clone is available through: Incyte Genomics, 4633 World  
Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or  
(314) 427-3222 FAX: (314) 427-3324. Web site:  
<http://www.incyte.com/reagents/catalog.jsp?page=clones/collaborative/index>  
Accession: F47499

## FEATURES

**Source**

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1083-3109"
/clone_1fb="Gm-r1083"
/note="The library Gm-r1083 is a sequence-driven, rerack-
ed set of 4,992 clones selected from cDNA libraries from
various tissues and stages of development of soybean. It
represents 1117 sequences from the progenitor library
Gm-cl1009 (from mature roots of 2 month old greenhouse
grown 'Williams' soybean plants); 820 sequences from the
progenitor library Gm-cl1013 (from 2 to 3 week old whole
plants of Williams); and 1055 sequences from library
Gm-cl1028 (from 'Superpod' plants whose seedlings were
inoculated with Bradyrhizobium japonicum, courtesy of Dr.
Gary Stacey). The 5' ends of the source clones from the
different progenitor libraries was used to select
singletons, or a representative of each contig, which were

```

RESULT	3
B1969937/c	
LOCUS	
DEFINITION	B1969937 G8630009911A07 Gm-r1083 Glycine max cDNA clone Gm-r1083-3109 3'
VERSION	B1969937
VERSION	B1969937.1 GI:16344342
KEYWORDS	EST.
SOURCE	Glycine max (soybean)
ORGANISM	Glycine max
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. 1 (bases) 1 to 7761

## REFERENCES

REFERENCE	1 (bases 1 to 776)
AUTHORS	Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Expelling, J., Rapp, C., Shoop, B., Pardini, J., Liu, L. and Lewin, E.
TITLE	A Functional Genomics Program for Soybean (NSF 9872565)
JOURNAL	Unpublished (1999)
COMMENT	Other PNs: AW810316 corresponding to Gm-c1028-4934 (5')

Contact: Vocklin, L.O., PI A Functional Genomics Program for Soybean (NSF 9872565)  
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics  
 University of Illinois  
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
 Tel: (217) 244-6147  
 Fax: (217) 333-4582  
 Email: l-vocklin@uiuc.edu  
 This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: <http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboration/index>  
 Seg primer: 5'-TTTTTTTTTTTTTTTTTTT/A/C/G/-3'.  
 Location/Qualifiers  
 1..776  
 /organism="Glycine max"  
 /mol\_type="rRNA"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1083-3109"  
 /clone\_1db="Gm-r1083"  
 /note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-c1009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-c1013 (from 2 to 3 week old whole plants of Williams); and 3095 sequences from library Gm-c1028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Rezel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.ahc.umn.edu/bioinformatics/nstsoy/>. Reracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.lie.niuc.edu/biocech/Keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."



Db 539 AGCTGATGTTGTTGAGACAGTTGAAAGTCCAGTCTCTCAGAGAGAAAGCTGTGTCAC 480  
 Qy 760 ACAAGATTCAAGGAGTTGTTGTTGTTATCCCTGTGTTCTTGAAGTCAATCTTTGTTG 819  
 Db 479 ACAAGATTCAAGGAGTTGTTGTTGTTATCCCTGTGTTCTTGAAGTCAATCTTTGTTG 420  
 Qy 820 ATGAAGTTGTTCAATATCAAGTATGAGAGCAATGAAAGTCCAAAGCTTCTGCGCTTA 879  
 Db 419 ATGAAGTTGTTCAATATCAAGTATGAGAGCAATGAAAGTCCAAAGCTTCTGCGCTTA 360  
 Qy 880 AAGAAGGCTATTTTGGGGAATATCTCCGAGCTGACAGTCTGCTCTTTTCAAGTTG 939  
 Db 359 AAGAAGGCTATTTTGGGGAATATCTCCGAGCTGACAGTCTGCTCTTTTCAAGTTG 300  
 Qy 940 CAAAAAGACCAAAATATGCGGGAAGCTTATTTGTCCTTTTCCAGCTTCCGAGGAGA 999  
 Db 299 CAAAAAGACCAAAATATGCGGGAAGCTTATTTGTCCTTTTCCAGCTTCCGAGGAGA 240  
 Qy 1000 GGTACCTGCTCCGCTGCTATTTGAGTCAAGTGAAGAGCAAGTGAAGTCACTTTG 1059  
 Db 239 GGTACCTGCTCCGCTGCTATTTGAGTCAAGTGAAGAGCAAGTGAAGTCACTTTG 180  
 Qy 1060 AGCCCTGATTCCTGCTTAAAGCTCTCACTAGTGAATTTTCTTGAATCTTGAACAGCT 1119  
 Db 179 AGCCCTGATTCCTGCTTAAAGCTCTCACTAGTGAATTTTCTTGAATCTTGAACAGCT 120  
 Qy 1120 TTAACATGATTTGTTAGAGTACTACTGTTTGAAGTCTGAGCTTAAATTAAGTCTGCTCC 1179  
 Db 119 TTAACATGATTTGTTAGAGTACTACTGTTTGAAGTCTGAGCTTAAATTAAGTCTGCTCC 60  
 Qy 1180 AAAAGACTAGTTTCTTATGATGCTCCCTGAGAGATATTTTGTGCTGCAATTAATA 1238  
 Db 59 AAAAGACTAGTTTCTTATGATGCTCCCTGAGAGATATTTTGTGCTGCAATTAATA 1

RESULT 5  
 BG648907 803 bp mRNA linear EST 24-APR-2001  
 LOCUS EST10526 HOGA Medicago truncatula cDNA clone pHOA-24A4 5' end,  
 DEFINITION  
 mRNA sequence.  
 BG648907.1 GI:13784019  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Medicago truncatula (Barrel medic)  
 Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabiales; Fabaceae; Papilionoideae; Trifoliales;  
 Medicago.  
 1 (bases 1 to 803)  
 Hahn,M.G., Ojansen-Reubs,T., Samac,D., Town,C.D., Van Aken,S.,  
 Ulteback,T., Cho,J. and Fraser,C.M.  
 ESTs from roots of Medicago truncatula treated with  
 oligogalacturonides of DP 6-20  
 Unpublished (2001)  
 Contact: Michael G. Hahn  
 Complex Carbohydrate Research Center  
 University of Georgia  
 220 Riverband Road, Athens, GA 30602-4712, USA  
 Tel: 706-542-4457  
 Fax: 706-542-4412  
 Email: hahnecrc@uga.edu  
 G392513e TIGR sequence name: MTRCRO2TK More information is  
 available at: www.medicago.org  
 Seg primer: SKmod (CTA gaa gta gat CC).  
 Location/Qualifiers  
 1..803  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="A17"  
 /db\_xref="taxon:3880"  
 /clone="pHOA-24A4"  
 /tissue\_type="3 day old seedling roots"

ORIGIN

/dev stase="24 hours after treatment in the dark at 26 C  
 with 0.5 mg/ml oligogalacturonides (DP 6-20) in the  
 presence of 100 ug/ml Gentamicin"  
 /lab\_host="XLOUP"  
 /clone\_id="HOGA"  
 /note="Vector: Bluescript SK-. Site 1: EcoRI, Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
 was directionally ligated into the Unizap XR vector from  
 Stratagene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in SOLR cells."

Query Match 47.8%; Score 651.4; DB 12; Length 803;  
 Best Local Similarity 91.3%; Pred. No. 5.3e-109;  
 Matches 724; Conservative 66; Mismatches 66; Indels 3; Gaps 3;

Qy 75 TAGTGAAGAT 134  
 Db 14 TAGTGAAGAT 72  
 Qy 135 ATTTGTTAAACCCATTTAGTATATCTAAATTAACCTTGGAGTGTGTTGCTCCGAGTT 194  
 Db 73 ATTTGTTAAACCCATTTAGTATATCTAAATTAACCTTGGAGTGTGTTGCTCCGAGTT 132  
 Qy 195 GCTGCTAAACTGAGATTGATGAGAGCAATGCTTATGTTGAAGAGACAGATTTGGTATAGT 254  
 Db 133 GCTGCTAAACTGAGATTGATGAGAGCAATGCTTATGTTGAAGAGACAGATTTGGTATAGT 192  
 Qy 255 ATGATTTGCTGATGAG 314  
 Db 193 ATGATTTGCTGATGAG 252  
 Qy 315 CCAACAGATGAT 374  
 Db 253 CCAACAGATGAT 312  
 Qy 375 CTGAT 434  
 Db 313 CTGAT 372  
 Qy 435 GGAGCTGATGTTGTTCTGATGAGATCTGCTTAAGAGATGAAGAGTGTCTGCAAGCT 494  
 Db 373 GGAGCTGATGTTGTTCTGATGAGATCTGCTTAAGAGATGAAGAGTGTCTGCAAGCT 432  
 Qy 495 GAAAGATATTTGCTTAAG 554  
 Db 433 GAAAGATATTTGCTTAAG 492  
 Qy 555 AATCCCAAGGTTCAATATGAAGCACTGCTCAGAGATATGAAGAGAGAGAGAGAGAGAG 614  
 Db 493 AATCCCAAGGTTCAATATGAAGCACTGCTCAGAGATATGAAGAGAGAGAGAGAGAGAG 552  
 Qy 615 ATGATGATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 674  
 Db 553 ATGATGATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 612  
 Qy 675 CTTAAAGAGCAAAATCCCAATTAAGCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 734  
 Db 613 CTTAAAGAGCAAAATCCCAATTAAGCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 672  
 Qy 735 CTCTCAG 794  
 Db 673 CTCTCAG 732  
 Qy 795 GGTGTTTGAAGTCAATCTTCTGATGAAGTTGTTCAATATCAAGTATGAAGAGATATA 854  
 Db 733 GGTGTTTGAAGTCAATCTTCTGATGAAGTTGTTCAATATCAAGTATGAAGAGATATA 790  
 Qy 855 GAAACTGCAAGAGC 867  
 Db 791 GAAACTGCAAGAGC 803

Mon May 24 08:18:50 2004

us-09-931-457a-30.rst

Page 6

RESULT 6  
AM311624 678 bp mRNA linear EST 02-DEC-2001  
LOCUS sg43b07.y1 Gm-cl025 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-cl025-1070 5', similar to TR:065747 O65747 CYSTEINE SYNTHASE,  
O-ACYTYL-L-SERINE, mRNA sequence.

ACCESSION AM311624  
VERSION AM311624.1 GI:6727270  
KEYWORDS EST  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max (soybean)

Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 678)  
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,  
Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Korn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800) 533-4363 or contact via email: csw@resgen.com  
Insert Length: 893 Std Error: 0.00  
High quality sequence stop: 475.  
Location/Qualifiers

FEATURES  
source  
1..678  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl025-1070"  
/tissue\_type="hypocotyl, 3 day seedlings"  
/lab\_host="DH10B"  
/clone\_id="Gm-cl025"  
/note="Vector: pTR73pac (pTR73, Pharmacia); Site 1: EcoRI;  
Site 2: NotI. This cDNA library was constructed from mRNA  
isolated from hypocotyl tissue of 3 day old seedlings.  
Complementary DNA was synthesized from mRNA using a poly  
(dT) primer with a NotI restriction site. EcoRI adapters  
were ligated to the blunt-ended cDNA fragments followed by  
digestion with EcoRI and NotI. The cDNA fragments were  
directionally cloned into the EcoRI-NotI restriction site  
of the pTR73-pac vector. The ligated cDNA fragments were  
transformed into DH10B host cells (Gibco BRL). This  
library was constructed by Dr. Randy Shoemaker."

# ORIGIN

Query Match 47.1%; Score 642; DB 10; Length 678;  
Best Local Similarity 98.2%; Pred. No. 3e-107;  
Matches 659; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 331 CTGGCATTTGATACCTTCAATGAGAGAGGAGGTTACAGCTATATTAATCAATG 390  
DB 8 CTGGCATTTGATACCTTCAATGAGAGAGGAGGTTACAGCTATATTAATCAATG 67  
QY 391 CTGCTTTATGAGCTTGAAGAGAGATCATTTATTAAGCTTTGAGCTGAGTTGTC 450  
DB 68 CTGCTTTATGAGCTTGAAGAGAGATCATTTATTAAGCTTTGAGCTGAGTTGTC 127  
QY 451 TGACAGATCTGCTAAGGAATGAAGGTCTGTTCAAGAGCTCAAGATATTTGGCTA 510

DB 128 TGACAGATCTGCTAAGGAATGAAGGTCTGTTCAAGAGCTCAAGATATTTGGCTA 187  
QY 511 AGAGCCCAATGCTTACATCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 570  
DB 188 AGAGCCCAATGCTTACATCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 247  
QY 571 ATGAACCACTGCTTCAAGATATGAGAAAGCTCCCATGAGAAATATGATGATTTGTT 630  
DB 248 ATGAACCACTGCTTCAAGATATGAGAAAGCTCCCATGAGAAATATGATGATTTGTT 307  
QY 631 CTGGATATGAGATCTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 690  
DB 308 CTGGATATGAGATCTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 367  
QY 691 CGAATATTAAGCTGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 750  
DB 368 CGAATATTAAGCTGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 427  
QY 751 CTGCTTCAACAGATTTCAAGAGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 810  
DB 428 CTGCTTCAACAGATTTCAAGAGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 487  
QY 811 ATCTTCTGATGATGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 870  
DB 488 ATCTTCTGATGATGATTTGCTTCAATGAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 547  
QY 871 TTGCGCTTAAGAGAGCTTATTTGAGAAATATCTTCCGAGCTTCAAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 930  
DB 548 TTGCGCTTAAGAGAGCTTATTTGAGAAATATCTTCCGAGCTTCAAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 607  
QY 931 TTGCGCTTAAGAGAGCTTATTTGAGAAATATCTTCCGAGCTTCAAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 989  
DB 608 TTGCGCTTAAGAGAGCTTATTTGAGAAATATCTTCCGAGCTTCAAGATTTGAAAACCTTCCCATCCCAAGTTCAAT 667  
QY 990 TTGCGGAGAGAG 1000  
DB 668 TTGCGGAGAGAG 678

RESULT 7  
CF513886 929 bp mRNA linear EST 09-SEP-2003  
LOCUS Cabud0007.IIIF.H10 Vitis vinifera cv. cabernet sauvignon (clone 8)  
DEFINITION Bud - CABUD Vitis vinifera cDNA clone Cabud0007.IIIF.H10 5', mRNA  
sequence.  
ACCESSION CF513886  
VERSION CF513886.1 GI:34545654  
KEYWORDS EST  
SOURCE Vitis vinifera  
ORGANISM Vitis vinifera  
REFERENCE Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.  
1 (bases 1 to 929)  
Goes da Silva, F., Iandolo, A., Llm, H., Baek, J., Jones, K. and  
Cook, D.  
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'  
berries at various developmental stages  
Unpublished (2003)  
Contact: Douglas Cook, PhD  
CARS Genome Facility  
UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drc@ucdavis.edu  
Seq primer: ACGTACCGGACATATGCC.  
Location/Qualifiers

FEATURES  
source  
1..929  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/cultiyar="Cabernet Sauvignon (clone 8)"

/db\_xref="taxon:29760"  
 /clone="Cabud0007\_11IF\_H10"  
 /sex="Hermaphrodite"  
 /dev\_stage="pre-bloom (10-11 days before bloom)"  
 /lab\_host="DHSalpha"  
 /clone\_lib="Vitis vinifera cv. cabernet sauvignon (clone  
 8) Bud - CABUD"  
 /note="Organ: Bud; Vector: pDNA; Site 1: SfiI; Site 2:  
 SfiI; CABUD is a cDNA library of Vitis vinifera cv.  
 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were  
 collected May 13, 2002 from pre-bloom plants (10-11 days  
 before bloom), pre-veraison. Sampled vines were located at  
 the University of California, Davis, Experimental  
 vineyard. cDNAs were made by oligo-dT priming and  
 directionally cloned. 5' and 3' adaptors were used in  
 cloning as follows:  
 5'-AAGCAGTGTATCAAGCCAGAGTGGCCATTACGCGCG-3' and  
 5'-ATTCTAGAGCCGAGGCGCGGCGGAGT-3' (30)NN-3'. Library was  
 constructed using the Clontech Creator SMART kit and  
 size-selected to contain the 0.5-3 kb size fraction."

Query Match 46.7%; Score 636.4; DB 14; Length 929;  
 Best Local Similarity 80.8%; Pred. No. 2.7e-106;  
 Matches 742; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

150 TTATGATATCTAAATTAACCTTGGAGTGTGTGTCGCCGGGTGCTGCTAACTGGG 209  
 1 TTATGATATCTAAATTAACCTTGGAGTGTGTGTCGCCGGGTGCTGCTAACTGGG 60  
 210 TTGATGAGCCATGCTCTAGTGTGAAGA CAGAGTGGGTATGATATGCTGATGCA 269  
 61 ATGATGAGCCCTTCTCTAGTGTGAAGA CAGAGTGGGTATGATATGATGCA 120  
 270 GAAAGAGAGGCTTATCAGCTGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 329  
 121 GAAAGAGAGGCTTATCAGCTGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 330 ACTGCAATGATGATGCTTCAAGCAGAGCAGAGGCTTACAGCTTACATTAATCAAG 389  
 181 ACTGCAATGATGATGCTTCAAGCAGAGCAGAGGCTTACAGCTTACATTAATCAAG 240  
 390 CTTGCTTCTATGATCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 449  
 241 CTTGCTTCTATGATCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 300  
 450 CTGACAGATCTCTGCTAAGGATGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509  
 301 CTGACAGATCTCTGCTAAGGATGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 510 AAGAGCCCATGCTCACTTCAACATTTGAAAACCTGCGAATCCCAAGTTCAT 569  
 361 AAGAGCCCATGCTCACTTCAACATTTGAAAACCTGCGAATCCCAAGTTCAT 420  
 570 TATGAAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629  
 421 TATGAAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 630 TCTGGGATAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689  
 481 TCTGGGATAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 690 CCGAATTAAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749  
 541 TCTGCACTTAAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 750 CTTGCTTCTATGATCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 809  
 601 CTTGCTTCTATGATCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 660  
 810 AATCTTCTGATGATGCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 869  
 661 AATCTTCTGATGATGCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 720

QY 870 CTTGCTTCTATGATGCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 929  
 DB 721 CTTGCTTCTATGATGCTTGAAGAGATCACTTCTATGCTTCTGAGCTGATGCTT 780  
 QY 930 TTTGAGATTTGCAAAAAGACAGAAAAGCCGGAAGCTTATGCTGCTGCTTTCACAGC 989  
 DB 781 ATTAAGTATGCAAAAAGACAGAAAAGCCGGAAGCTTATGCTGCTGCTTTCACAGC 840  
 QY 990 TTGCGGAGAGCTTCT 1049  
 DB 841 TTGAGAGAGCTTCT 900  
 QY 1050 ATGACTTTTGAGCCCTTGA 1067  
 DB 901 ATGCTTTTGAGCCCTTGA 918

RESULT 8  
 CA801673  
 LOCUS  
 DEFINITION  
 CA801673  
 632 bp mRNA linear EST 05-DEC-2002  
 sat16d04.y1 Gm-cl036 glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl036-13975 5' similar to TR:023733 023733 O-ACETYL SERINE (THIOI)  
 LVSAB.; mRNA sequence.  
 ACCESSION  
 CA801673  
 VERSION  
 CA801673.1 GI:26058759  
 EST.  
 SOURCE  
 Glycine max (soybean)  
 ORGANISM  
 Glycine max (soybean)

REFERENCE  
 AUTHORS  
 1 (bases 1 to 632)  
 Shoemaker, R., Kaim, P., Vodkin, L., Espelding, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,  
 Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritters, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterson, R., and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 TITLE  
 Public Soybean EST Project  
 JOURNAL  
 COMMENT  
 Contact: Shoemaker R./Public Soybean EST project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@waterson.wustl.edu

FEATURES  
 source  
 High quality sequence stop: 448.  
 Location/Qualifiers  
 1..632  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl036-13975"  
 /tissue\_type="somatic embryos cultured on MSD 20"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This  
 cDNA library was constructed from mRNA isolated from  
 somatic embryos (age ranging from 2 months to 9 months)  
 cultured on MSD 20. The library was prepared using the  
 Life Technologies puferScript cDNA library construction  
 kit. Complementary DNA was synthesized from mRNA using a  
 poly (dAT) sequence with a NotI restriction site. SalI  
 linkers adapters were ligated to the blunt-ended cDNA  
 fragments followed by NotI digestion. The cDNA fragments







```

QY 591 AATGGAAGGCTCCGATGGGAAATATGATGATTTGTTTCTGGATAGGCACTGGTGT 650
DB 488 ATCTGAGAGGACGACGCTGGGAAAGTGAATGTTTGTCTCTGGGATAGGCACTGGAGGT 547
QY 651 ACAATAGAGGTGCTGGAATATCTTAAAGACAGATCCGATATTAAGCTGATTTGT 710
DB 548 ACCATTAAGGTGACGGAAGTTCCTCAAAAGCAAAATTTGACTTAAGCTGATTTGT 607
QY 711 GTGGAACAGTTGAATGCTGCTCTGAGAGGAAGGCTGCTGCAACAGATTTGA 770
DB 608 GTAGAACCTGTGGAAGGCTGCTGCTGAGAGGAAGGCTGCTGCAACAGATTTGA 667
QY 771 GGGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
DB 668 GGAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
QY 831 CAATATATCAAGTATGATGAGCAATAGAACTGCAAGCTTCTGCTTAAAGAGGCTTA 890
DB 728 CAAGTTTCAAGTATGAGCAATAGAACTGCAAGCTTCTGCTTAAAGAGGCTTA 787
QY 891 TTTTGGGAATATCTTCCGAGCTGAGCTGCTGCTGCTTCAAGTTGCAAAAGACCA 950
DB 788 CTGGTTGGGATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
QY 951 GAAATATGCTGGAAGCTTAT-TGTTCCGCTTTTCCAGCTTCCGAGAGGCTTACCTGTC 1009
DB 848 GAAATATGCTGGAAGCTTAT-TGTTCCGCTTTTCCAGCTTCCGAGAGGCTTACCTGTC 907
QY 1010 CTC 1012
DB 908 CTC 910

RESULT 10
LOCUS BF003458 724 bp mRNA linear EST 06-OCT-2000
DEFINITION EST131956 KVI Medicago truncatula cDNA clone pKVI-5N8, mRNA
ACCESSION BF003458
VERSION BF003458
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula (barrel medic); Eubryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 724)
AUTHORS Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,
Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.B. and
Fraser, C.M.
TITLE ESTs from roots of Medicago truncatula 24 hours after inoculation
with Sinorhizobium meliloti
JOURNAL Unpublished (1999)
COMMENT Contact: Vandenbosch K
Department of Plant Biology
University of Minnesota
220 BiocSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cms.umn.edu
Texas A&M University name: T267268e TIGR sequence name: MT1A76TK
More information is available at: http://chrysis.tamu.edu/medicago
Seq primer: SKmod (CTA GAA CTA gtc gat CC).
FEATURES
source
1..724
location/Qualifiers
1..724
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKVI-5N8"
/tissue_type="Seedling roots"
/dev_stage="24 hours post-inoculation with Sinorhizobium

```

```

meliloti"
/lab host="E. coli strain XL0R"
/clone_lib="KVI"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0R cells."

ORIGIN
Query Match 46.0%; Score 627.2; DB 10; Length 724;
Best Local Similarity 91.9%; Pred. No. 1.4e-104;
Matches 662; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 96 GTTGAAGGTCGGAATTCGCAAGATGTTACGGAATTGATGTTAAACCCCAATTGTA 155
DB 5 GGTGAAGGTCGGAATTCGCAAGATGTTACGGAATTGATGTTAAACCCCAATTGTA 64
QY 156 TATCTAAATTAATCTGCGATGTTGTTGCTCCCGGTTGCTGCTTAACTGAGTTATG 215
DB 65 TATCTAAATTAATCTGCGATGTTGTTGCTCCCGGTTGCTGCTTAACTGAGTTATG 124
QY 216 GAGCCATCTTATGTTGAAGACAGATTTGGTATGATGATTTGCTGATGACGAAGAG 275
DB 125 GAGCCATCTTATGTTGAAGACAGATTTGGTATGATGATTTGCTGATGACGAAGAG 184
QY 276 AAGGACCTTATCAACCTGGAAGAGTGTCTCTCATTTGAGCCACAGAGTGTATATGCG 335
DB 185 AAGGACCTTATCAACCTGGAAGAGTGTCTCTCATTTGAGCCACAGAGTGTATATGCG 244
QY 336 ATTGATTAAGCTTCTATGAGCAGAGCCAGGGTTTCAAGCTCATATTAATGCTGCT 395
DB 245 ATTGATTAAGCTTCTATGAGCAGAGCCAGGGTTTCAAGCTCATATTAATGCTGCT 304
QY 396 TCTATGAGCTTGAAGAGATCATTTATTTAGCTTTGAGCTGAGTTGTTGAGCA 455
DB 305 TCTATGAGCTTGAAGAGATCATTTATTTAGCTTTGAGCTGAGTTGTTGAGCA 364
QY 456 GATCTGCTTAAGGATGAAGGCTGCTGTTCAAGAGCTGAAGATTTGCTTAAGCG 515
DB 365 GATCTGCTTAAGGATGAAGGCTGCTGTTCAAGAGCTGAAGATTTGCTTAAGCG 424
QY 516 CCACATGCTTCAATCTTCAACATTTGAACCTGCAATCCCAAGTTCAATATGA 575
DB 425 CCACATGCTTCAATCTTCAACATTTGAACCTGCAATCCCAAGTTCAATATGA 484
QY 576 ACCATGCTTCAAGATATGAAGGCTCCGATGGAATTTGATGCAATTTGTTCTGGG 635
DB 485 ACCATGCTTCAAGATATGAAGGCTCCGATGGAATTTGATGCAATTTGTTCTGGG 544
QY 636 ATGACATTTGTTGATCAATTAACAGTGTCTGGAATTTCTTAAGACCAATCCGAT 695
DB 545 ATGACATTTGTTGATCAATTAACAGTGTCTGGAATTTCTTAAGACCAATCCGAT 604
QY 696 ATTAAGCTGATTTGTTGATCAATTAACAGTGTCTGGAATTTCTTAAGACCAATCCGAT 755
DB 605 ATTAAGCTGATTTGTTGATCAATTAACAGTGTCTGGAATTTCTTAAGACCAATCCGAT 664
QY 756 CCACATGATTTCAAGGATTTGTTGTTTATCTCTGTTGTTTGAAGTCAATCTT 815
DB 665 CCTCAAGATTTCAAGGATTTGTTGTTTATCTCTGTTGTTTGAAGTCAATCTT 724

RESULT 11
BG595822 791 bp mRNA linear EST 07-MAR-2003
LOCUS BG595822
DEFINITION EST194500 cSTS solanum tuberosum cDNA clone cSTS1N11 5' sequence,
ACCESSION BG595822
VERSION BG595822
KEYWORDS BG595822.1 GI:13613962
EST.

```

SOURCE  
ORGANISM  
Solanum tuberosum (potato)  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 791)  
REFERENCE  
AUTHORS  
van der Hoeven R., Bezzerides J., Sun H., Cho J., Chieningo A.,  
Bougti O., Buell C.R., Romling C., Tanksley S. and Baker B.  
Generations of ESTs from sprouting potato eyes  
Unpublished (2000)  
CONTACT: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr. Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: M13F-R.  
FEATURES  
Source  
1..791  
location/Qualifiers  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cST511N11"  
/tissue\_type="sprouting eyes from tubers"  
/dev\_stage="12-14 weeks post harvest"  
/lab\_host="SOLR"  
/clone\_lib="cSTS"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XbaI; Various sizes of sprouting eyes (2mm to 15mm) were  
taken from tubers. The tubers were incubated at 26C in the  
dark for 2-3 weeks prior to sprouting. The eyes were  
frozen in liquid nitrogen immediately upon removal from  
tubers."

ORIGIN  
Query Match 46.0%; Score 626.2; DB 12; Length 791;  
Best Local Similarity 89.7%; Pred. No. 2.1e-104;  
Matches 695; Conservative 0; Mismatches 78; Indels 2; Gaps 2;  
QY 51 TCCTGTTCTCAAGTCAGCTTTGAGTACAGATTAATGAGCTGTGAAGTCGGA 110  
DB 12 TCCATTTCTCAACACACACAAAGTTAGTGAAGCTATACAGCGGCTGAAA-GTCTGA 70  
QY 111 ATTGCCAAGATGTACCGAATTGATGTAATGTAATGTAATGTAATGTAATGTA 170  
DB 71 ATTGCTAAGATGTTACGAATGATGTAATGTAATGTAATGTAATGTAATGTA 130  
QY 171 GCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 230  
DB 131 GCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 190  
QY 231 GTGAAGACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 290  
DB 191 GTGAAGACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 250  
QY 291 CCTGGAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 350  
DB 251 CCTGGAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 310  
QY 351 ATGGACAGAGCGAGGTTTACAGCTCAATTAATTAATTAATTAATTAATTAATTA 410  
DB 311 ATGGACAGAGCGAGGTTTACAGCTCAATTAATTAATTAATTAATTAATTAATTA 370  
QY 411 AGAAGATCAATTTCTTATGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 470  
DB 371 AGAAGATCAATTTCTTATGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 430  
QY 471 ATGAAGAGTGTCTCAGAGGCTGAGAGATATTGCTTAAGACCCCAATGCTTACATA 530  
DB 431 ATGAAGAGTGTCTCAGAGGCTGAGAGATATTGCTTAAAGACCCCAATGCTTACATA 490

QY 531 CTTCACAAATTGAAAACCTGCCAATCCCAAGTTCAATTATGAAAACCACTGTCGAG 590  
DB 491 CTTCACAAATTGAAAACCTGCCAATCCCAAGTTCAATTATGAAAACCACTGTCGAG 550  
QY 591 ATATGAAAAGCTCCGATGAGAAAATTGATGATTTGTTTTCGGATATAGCACTGTGT 650  
DB 551 ATATGAAAAGCAAGATGAGAAAATTGATGATTTGTTTTCGGATATAGCACTGTGT 610  
QY 651 ACATTAACAGGTGCTGAGAAAATATCTTAAAGACAGAAATCCCAATTAAGCTGA-TTGG 709  
DB 611 ACATTAACAGGTGCTGAGAAAATATCTTAAAGACAGAAATCCCAATTAAGCTGA-TTGG 670  
QY 710 TGTGGAACCAAGTTGAAAGTCCAGTCTCTCAGAGAAAAGCTGTCCACAAATTTCA 769  
DB 671 TGTGGAACCAAGTTGAAAGTCCAGTCTCTCAGAGAAAAGCTGTCCACAAATTTCA 730  
QY 770 AGGATTTGATGCTGCTTTTATCCCTGAGTCTTGAAGTCAATCTTGAATGA 824  
DB 731 AGGATTTGATGCTGCTTTTATCCCTGAGTCTTGAAGTCAATCTTGAATGA 785

RESULT 12  
CK260011 994 bp mRNA linear EST 12-DEC-2003  
LOCUS B5706089 potato abiotic stress cDNA library Solanum tuberosum cDNA  
DEFINITION  
Accession  
CK260011 GI:39816989  
VERSION  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Solanum tuberosum (potato)  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 994)  
REFERENCE  
AUTHORS  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
Generation of ESTs from abiotic stressed potato tissue  
Unpublished (2003)  
OTHER ESTS: B5706088 B5706090 B5706091  
CONTACT: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr. Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
Source  
1..994  
location/Qualifiers  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB228"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DRI08-T02A"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat





Page 13

RESULT 15					
BO630428					
LOCUS					
DEFINITION	BO630428	613 bp	mRNA	linear	EST 21-OCT-2007
	sa007c05.y1 Gm-cl045	glycine max	cdna clone	SOYBEAN	CLONE ID:
	Gm-cl045-1969 5'	similar to	TR:081154	081154	CYSTEINE SYNTHASE. ;
	mRNA sequence.				
ACCESSION	BO630428				
VERSION	BO630428.1	GI:21678077			
KEYWORDS	EST.				
SOURCE	glycine max (soybean)				
ORGANISM	glycine max				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 613)	Shoemaker, R., Keim, P., Vodka, L., Erpelndg, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kuchta, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, J., Pearson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, V., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	Contact: Shoemaker R/Public Soybean EST Project

FEATURES	Location/Qualifiers
source	1. .613

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl045-3969"
/tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
/lab_host="DH10B"
/clone_lib="Gm-cl045"
/notes="Vector: pBluescriptII SK; Site_1: EcoRI, Site_2: XhoI; This cDNA library was constructed from mRNA isolated from etiolated hypocotyl tissue of 9-10 day old seedlings of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) primer with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and XhoI. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Schenaker."

```

```

Query Match= 44.5%; Score 606; DB 13; Length 613;
Best Local Similarity 100.0%; Pred. No. 1..1e-100;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 9 GTTCTAGATACCCAGATGCTGTTGCTTAGTGTGTCAGTCATTCCTGTTCTCTCAAGTCAA 68
DB 8 GTTCTAGATACCCAGATGCTGTTGCTTAGTGTGTCAGTCATTCCTGTTCTCTCAAGTCAA 67
QY 69 GCTTGTGATGAGCAGATTAATGAGCTGTGAAGAGGTCCGGAATTCGCAAGAGTGAAC 128
DB 68 GCTTGTGATGAGCAGATTAATGAGCTGTGAAGAGGTCCGGAATTCGCAAGAGTGAAC 127
QY 129 GAATGATTTGGTAAACCCCATTAAGTATATCTAAATAAACCTTGCGAGTGTGTGTC 188
DB 128 GAATGATTTGGTAAACCCCATTAAGTATATCTAAATAAACCTTGCGAGTGTGTGTC 187
QY 169 CGGTTGCTGCTAAACCTGAGTTGATGAGCCATGCTAGTGTGAAGACAGATTTGAG 248
DB 168 CGGTTGCTGCTAAACCTGAGTTGATGAGCCATGCTAGTGTGAAGACAGATTTGAG 247
QY 249 TATAGTATGATTTGCTGATGACAGAGAGAGGAGCTTATCACAACCTTGAAAGAGTGTCTC 308
DB 248 TATAGTATGATTTGCTGATGACAGAGAGAGGAGCTTATCACAACCTTGAAAGAGTGTCTC 307
QY 309 ATTGAGCCAAACAGGTGTATCTGCACTTGAATTGACCTTCAATGACACAGCAAGGAGT 368
DB 308 ATTGAGCCAAACAGGTGTATCTGCACTTGAATTGACCTTCAATGACACAGCAAGGAGT 367
QY 369 TACAAGCTCATATATTACAATGCTGCTCTTCTATGAGTCTTGAGAGAGAAATCAATTATTA 428
DB 368 TACAAGCTCATATATTACAATGCTGCTCTTCTATGAGTCTTGAGAGAGAAATCAATTATTA 427
QY 429 GCTTTTGAAGCTGAATTTGGTTCTGACAGATTCCTGCTTAAGAGAAATGAAGAGTGTGAC 488
DB 428 GCTTTTGAAGCTGAATTTGGTTCTGACAGATTCCTGCTTAAGAGAAATGAAGAGTGTGAC 487
QY 489 AAGGCTGAAGAGATTTGGCTTAAGACGCCCAATGCTACATATCTTCAACATTTGAAAC 548
DB 488 AAGGCTGAAGAGATTTGGCTTAAGACGCCCAATGCTACATATCTTCAACATTTGAAAC 547
QY 549 CTTGCGCATTCCTCAAGCTCATTAATGAACCACTGTGCCAGATTAATGAAGAGTCCGAT 608
DB 548 CTTGCGCATTCCTCAAGCTCATTAATGAACCACTGTGCCAGATTAATGAAGAGTCCGAT 607
QY 609 GGGAAA 614
DB 608 GGGAAA 613

Search completed: May 21, 2004, 18:59:53
Job time : 3965 secs

```

Search completed: May 21, 2004, 18:59:53  
Job time : 3965 secs

ORIGIN



Mon May 24 08:18:49 2004

us-09-931-457a-30.rni

**Page 2**

```

Db      301  CCTCATCAATGAGGCTTGAGAGAAAATTCATTCTTTAGCATTTAGTGGCGGAGGTTAC 360
Oy      450  CTGACAGATCTCTGCTAAGGGAATGAAGATGCTGTTCGAAGGCTGAGAGATATATGGCT 509
Db      361  CTCACAGATCTCTGCTAAGGGAATGAAGATGCTGTTCGAAGGCTGAGAGATATATGGCT 420
Oy      510  AAGAGCGCCCAATGCTCTACATATCTTCAACATATTTGAAAAACCTGCCAATCCCAAGGTTAT 569
Db      421  AAAATATCCAGATATGATCATGTCTAGAACAGTTCAAAAATCTCTTCAAAACCGCAAACTCAT 480
Oy      570  TATGAAACCATCTGCTCCAGATATAGAAAGGCTCCGATGGGAAATTTGATGCAATTTGTT 629
Db      481  TATGAAACCATCTGCTCCAGATATAGAAAGGCTCCGATGGGAAATTTGATGCAATTTGTT 540
Oy      630  TCTGGGATAGGCACTGCTGTGTCACATTAACAGGCTGTGAAAAATATCTTAAAGACAGAT 689
Db      541  GCCGCTGTGGAATCGTGGAACGCTTCCGATCAGAGAAATTCCTCAAGAGAAAGAT 600
Oy      690  CCGAATATAAAGCTGATTTGTTGTGGAACCAAGTTGAAGTCCAGTCTCTCAGAGAGAAAG 749
Db      601  AAAGCTTTAAAGTTATGATGTGTGAACCTTACAGAAATGTCCGTTATATAGTGAAGCAAA 660
Oy      750  CCGTGTCCACCAACATTCAGAGGATGGTGTGCTGATTTATCCCTGATGTCTTGTGAGATC 809
Db      661  CCGGTTACATTTGATCCAGATATTTGGGCTGACATCCAGACAAATTTGATTTTC 720
Oy      810  AATCTTGTGATGAGTTGTTCAAAATATCAAGTATGAGACATTAAGAACTGCAAAAGCTT 869
Db      721  AAGCTTCTGTGATGAGTATCCAGTGCACAGTGTGGAAGCAATTAAGAAAGCAAACTT 780
Oy      870  CTTGGGCTTTAAGAGAGGCTCTATTTTGTGGAATATCTTCGGAGCTGCAGCTCTGCT 929
Db      781  CTTGGGCTTTAAGAGAGGCTCTATTTTGTGGAATATCTTCGGAGCTGCAGCTCTGCT 840
Oy      930  TTTGAGATTTGCAAAAAGACCAAGAAATGCCGGGAAGCTTATTTGTTGCCGTTTCCAGC 989
Db      841  ATAAAGGTGGCAAAACCGGCTCAGAAAAACCGGCAAACTATGTTGTGATTTTCCCTAGC 900
Oy      990  TTGGGGAGAGGTACTGTCCTCCGATTTAGTGTAGTAGACGGGAGGTGAAGC 1045
Db      901  GGTGGGAAAGCTTACTTATGACTCTCATTTGTTGATGATGACATGACATGAAAGAGAT 960
Oy      1050  ATGACTTTGA 1060
Db      961  TTGCCAATTC 971

RESULT 2
US-09-596-002-39
/ Sequence 39, Application US/09596002
/ Patent No. 6632636
/ GENERAL INFORMATION:
/ APPLICANT: Lagace, Robert, E.
/ APPLICANT: Patterson, Chandra
/ APPLICANT: Berg, Kim, L.
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
/ FILE REFERENCE: PM-0008-4 US
/ CURRENT APPLICATION NUMBER: US/09/596,002
/ PRIOR FILING DATE: 2000-06-16
/ PRIOR APPLICATION NUMBER: 60/140,121
/ PRIOR FILING DATE: 1999-06-18
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PERL Program
/ SEQ ID NO 39
/ LENGTH: 100848
/ TYPE: DNA
/ ORGANISM: M. catarrhalis
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte template ID No. 6632636 39
/ PUBLICATION INFORMATION:
/ US-09-596-002-39

```

Query Match	20.0%;	Score 273;	DB 4;	Length 10084;
Best Local Similarity	57.7%;	Pred. No. 3.1e-67;		
Matches 527;	Conservative 0;	Mismatches 380;	Indels 6;	Gaps 2;

OY	111	ATTGCCAAGATGTTACGGAAATTGAATTGTATAAACCCCATTAAGTATATCTTAATAACTT	170
Db	17953	ATTGCCAAAAAGTGTGACTACCTTATTCGTATAAACCCCGCTTGTAATCAATGCCTC	18012
OY	171	GCGAGATGTTGTGTACC CGGGTGTGCTGTAACTGAGTTGATGAGCATGCTCTAGT	230
Db	18013	ACCGAAGCTTATCTCTCTGTGTGTGTGTGTCAAGCTTAGTATGCCAATCACCAAGTTGG	18072
OY	231	GTGAGAACAAGATTGGGTATAGTATGATTGCTGTATGCAAGAAAGAAAGGAACTTATACA	290
Db	18073	GTCAAGAAATGCACTTTCCTCTCTATGATTGACATGCGAAAATTCAGGGCTTTTGA--	18130
OY	291	CCTGAAAAGATGTCTCTATTGAGCAACAAGATGTGAATACGTGGCATTTGATTAACCTTC	350
Db	18131	-CCAAGACACACACATGTTGAGGGAGCAAGGGAAATACAGGCACTTGGCTTGCAAG	18189
OY	351	ATGGCAACAGCCAGAGGGTTACAAGCTCAATTAATTAATGCTCTGTCTTATGAGTCTTAG	410
Db	18190	GTGGCAGCGGCAAGAGSTTACCGCTTGGATTACCAATGCTGGAAGCAATGAGCCCTTAG	18249
OY	411	AGAAGATCAATCTATTAGCTTTTGGAGCTGAGTGGTGTCTGACAGATCTCTCTAAGGA	470
Db	18250	CGTGTGCGTGTCTGTGTGTGTATGAGGGCAAGCTGTGTGTATACCGCAAGCGCTGAGGGC	18309
OY	471	ATGAAAAGTGTCTGTCAAGAGCTGAAGAGATATTGGCTTAAGACGCCCAAATGCCATACMA	530
Db	18310	ATGAGTGGGGCGAATATCCCAAGCAATGAGCTTCTGCCCAAGACGGCTAT--TTTAG	18366
OY	531	CTTCAACAATTTGAAAAACCTCGCCCAATCCCAAGTTCATTATGAAACCACTGGTCCAGAG	590
Db	18367	CAAGCGCAATTTGATTAATTTAGCAATTCACAATTCACCGTATCAACACACGCCGAAGG	18426
OY	591	ATTATGAAAAGGCTCCCATGAGAAATATTATGATTTGTTCTGGAGTATAGGACTGTGTGT	650
Db	18427	ATTATGGCGGATACAAATGATGTAGAGGATATTTTGTCTCAAGTGTGGSTACAGGTGCT	18486
OY	651	ACATATAACAGTGTGTAATAATCTTAAGACAGAAATCCGATATAAAGCTGATTTGCT	710
Db	18487	ACAGTACAAGTGTGTGTGTAGACACTCAAAAGCAAAAAGAACCATTCATAATTTGTGCA	18546
OY	711	GTGGAACCAATTGAAAGTCCAGTGTCTCAAGAGGAAAGCTGTGTCCACACAAGATTCAA	770
Db	18547	GTTTGAGCAATTGGCAATCCCTGTCTTATAGAGGAGGAAAAAGGCCGCAACCAATTCAA	18606
OY	771	GGGATTGTGCTGGTTTTTATCCCTGTGTCTTGAAGTCAATCTTTCTTGATGAAGTTGT	830
Db	18607	GGCATTGTGTCAAGGTATTATCCCATGAGTACTCAATACGATATCATGATGAGGTATC	18666
OY	831	CAAAATCAAGTATATAGCAATAGAAATCGCAAAAGCTTCTTGCGCTTAAABAAGGCCCA	890
Db	18667	AAAGTGTCTATATAAGATGCTTTAGTGTGGCAACGATATGTGCAAGCCCAAGAAAGGCTTG	18726
OY	891	TTTGTGGAAATATCTTCCGAGCTGCAGCTGTCTCTTTTCAATTGCCAAAAAGACCA	950
Db	18727	CTTGTGTGATCTCATCAAGTGTGCGAGTTTGGCGGCACTTGAAATGTGCGGCAAGAGA	18786
OY	951	GAATAATGCCGGAAGCTTATTTGTGCCCCCTTTTCCAGCTGTGGGGAGAGGTACCGTGTCC	1010
Db	18787	GAATAAGCAAGTAAAGCTTATGTCGTATCAATCCATCATCAAGGTAGCGCTTATTTGTCA	18846
OY	1011	TTCGTGCTATTTG	1023
Db	18847	ACAGCAATTATTG	18859



GENERAL INFORMATION:  
 APPLICANT: GARY BRETON  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709.1002-001  
 CURRENT APPLICATION NUMBER: US/09/543,681A  
 CURRENT FILING DATE: 2000-04-05  
 PRIOR APPLICATION NUMBER: US 60/128,706  
 PRIOR FILING DATE: 1999-04-09  
 NUMBER OF SEQ ID NOS: 8344  
 SEQ ID NO 3561  
 LENGTH: 1008  
 TYPE: DNA  
 ORGANISM: Proteus mirabilis  
 US-09-543-681A-3561

Query Match 18.8%; Score 255.4; DB 4; Length 1008;  
 Best Local Similarity 58.3%; Pred. No. 2.8e-63;  
 Matches 561; Conservative 0; Mismatches 366; Indels 36; Gaps 5;

85 ATATATAGCTGTGAAAGTCCGGAATTCGCAAGATTTACGGAATTTGATGTTAA 144  
 38 ATGTACCGAGATTAAGGATTAATTTGGAATTAATCTGCAAACTATGTCCTA 97  
 145 CCCCATATATATCTAAATTAATCTGCGATGTGTGTGCCCCGCTGCTGCTAAC 204  
 98 CCCCCCTATGCTTTAAATCAATTCG-----GTAATGCAATATCTAGCTTAAG 148  
 205 TGAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 264  
 149 TAGAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 208  
 265 ATGCAAG 324  
 209 ATGCAAG 265  
 335 GTATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 384  
 266 GTATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325  
 385 CATATCTGCTCTATGATCTGATGATGATGATGATGATGATGATGATGATGAT 444  
 326 CATATCTGCTCTATGATCTGATGATGATGATGATGATGATGATGATGATGAT 385  
 445 TGGTCTGACAGATCTGCTGATGATGATGATGATGATGATGATGATGATGATG 504  
 386 TGTGTGTAATGAG 445  
 505 TGGCTATGAG 561  
 446 GTGATAG 505  
 562 AGGTCTATATGAAACCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621  
 506 AATTCATATGAAACCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 565  
 622 CATTTGTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 681  
 566 TGTGTTATCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 625  
 682 A-----GAGAGATCCAGATGATGATGATGATGATGATGATGATGATGATG 738  
 626 ATACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 685  
 739 CAG 780  
 686 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 745  
 781 CTGCTTTATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 746 CAGGTTTATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 805  
 841 GTGATGAG 900

806 ACGAG 865  
 901 TATTTCCGAG 960  
 866 TTTTCATCCGAG 925  
 961 GGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 926 ATAG 985  
 1021 TTG 1023  
 986 TTG 988

RESULT 4  
 US-09-328-352-1324  
 Sequence 1324, Application US/09328352  
 Patent No. 6562958  
 GENERAL INFORMATION:  
 APPLICANT: Gary L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 CURRENT FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252  
 SEQ ID NO 1324  
 LENGTH: 1008  
 TYPE: DNA  
 ORGANISM: Acinetobacter baumannii  
 US-09-328-352-1324

Query Match 18.4%; Score 251.2; DB 4; Length 1008;  
 Best Local Similarity 58.3%; Pred. No. 4.4e-62;  
 Matches 549; Conservative 0; Mismatches 363; Indels 30; Gaps 5;

107 CGAATTCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 166  
 54 CGAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 110  
 167 ACTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 226  
 111 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 170  
 227 TAGTGTGAG 286  
 171 TTTGTGTAATGAG 230  
 287 CAGCTGGAAG 346  
 231 GAAAG 287  
 347 CTTCATGAG 406  
 288 GTTGTGAG 347  
 407 TGAAG 466  
 348 TGAAG 407  
 467 GGAAG 526  
 408 GGAAG 467  
 527 ---CATCTTCAAGATTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583  
 468 TTTCTTGCAG 527  
 584 TCCAGAGATGAG 643  
 528 ACCAGAGATGAG 587



Mon May 24 08:18:49 2004

us-09-931-457a-30.rnt

Page 4

QY 644 TGGTGGTACATATACAGTGTGGAATAATCTTA---AAGACAGATCCGAAATPAA 700  
DB 588 AGAGGTACATTAACGGTATTTCTCGTATTTTGAACAAGTTCAAGATTAAGCATGTA 647  
QY 701 GCTGATTTGTGTGACCAAGTTGAAAGTCCAGTCTCTGAGAGGAAGCTGT----- 755  
DB 648 CTGATGACGATTTGACCTGTGAGTCTCCATTTATTAACAACAACAAAATGTTAGAA 707  
QY 756 -----CCACAGAGATTCAGGATTTGAGTGTGTTTATCCCTGGTCTT 802  
DB 708 CATTAACCCGACCCGATTAATTCAGGATTTGAGCAAAATTTATTCAAAAGATCT 767  
QY 803 GGAAGTCAATCTTCTGATGAAGTTTCAAAATCAAGTGAAGCAATAGAACTGC 862  
DB 768 TGATTTAGATTTGTTGATGAAGTCTTACAGTGAAGTGAAGAACAAATTCATGAGC 827  
QY 863 AAAGCTTGTGCGCTTAAAGAGCCCTATTGTGGAATATCTCCGAGCTGAGCTGC 922  
DB 828 GAGAAAGTGTCCACCCAGAGGATTTTATGAGGATTTTCAAGTGTCTGATGCG 887  
QY 923 TGCTGCTTTTCAATTTGCAAAAAGACAGAAATGCGGAGGCTTATTTGCGGTTT 982  
DB 888 AGCAGTGGCTAAATTTGAGAACCTCTGAAATGCAAGTAAACAATGTTGTTATCT 947  
QY 983 TCCAGCTTCCGGGAGAGTACCTGCTCCGCTGCTATTTGA 1024  
DB 948 TCCAGATAGTGGGAGCCCTATTTATCTTCTGTTTATTTGA 989

RESULT 5

US-08-956-171E-148  
Sequence 148, Application US/08956171E  
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESSES:

ADDRESS: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956.171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009.861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,286

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 148:

SEQUENCE CHARACTERISTICS:

LENGTH: 6115 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-08-956-171E-148

Query Match 18.4%; Score 251.2; DB 4; Length 6115;  
Best Local Similarity 58.7%; Pred. No. 1.2e-61;  
Matches 493; Conservative 0; Mismatches 338; Indels 9; Gaps 3;

QY 121 ATGTTACGAAATTTGATTTGTTAAACCCATTTATCTTAATAAATCTGCGATGTT 180  
DB 4760 ATTTTACTGAATTTATTTGGCGGTGACCCGGTGTCAATTTGAGAAATGATGATGACA 4819  
QY 181 GTGTGTCGGGTGCTGCTAAATCTGAGTTGATGAGGCAATGCTCTAGTGAAGACA 240  
DB 4820 ATGACGCAAGTGTATATGTAATAATGGAATTCAAAATTCAGGTGTTCTGTAAAGATA 4879  
QY 241 GGAATGGTATAGTATGATTTGCTGATGACAGAAAGAGGAGCTTATCACACTGAAAGA 300  
DB 4880 GAATTCCTTTAGCAATGATTTGAAAAGAGAGAGGAGCAAAATTAACCTGGCGA-- 4937  
QY 301 GTGTCTCATTTAGCCCAACAGTGTATATCTGCAATTTGATTTAGCTTCAATGCGAG 360  
DB 4938 -TACAAATTTAGAACCAACAGAGGTATATACAGTATCGTTAGCATTTGATGTCTG 4996  
QY 361 CCGAGGGTTTACAAGCTCATTAATTAACATGCTGCTCTATGAGTCTTGAGAGAGATCA 420  
DB 4997 CTAAAGATATTAAGAGATTTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5056  
QY 421 TTCTATTAGCTTTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB 5057 TATTTAAAGCATGCGGTGCGGAAATGTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTG 5116  
QY 481 CTGTTCAGAGGCTGAGAGATTTGCTTAAAGCCCAATGCTCTCATTAATTTCAACAT 540  
DB 5117 CAATTAATAAGTATTAAGA--ATTAAGAAAGAAATGCTTACTTGAGCCCAACAT 5173  
QY 541 TTGAAAACCTGCAATCCCAAGCTTCAATTATGAACCACTGCTCAGAGATTTGGAAG 600  
DB 5174 TTGAAAACCTGCAACCTGCAAGCTTCAATGATTAATCAAGTCTCTGAGTTATTTCAAC 5233  
QY 601 GCTCCAGTGGGAAATTT---GATGCTTTGTTTCTGAGATGAGCACTGAGTACATAA 657  
DB 5234 AATTTAAGGAAATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5293  
QY 658 CAGTGTGTAATAATCTTTAAAGACAGAAATCCGAATTAAGCTGATTTGTGTGAAC 717  
DB 5294 CTGTGTAGTAAAGTTCTGAAAAGAAATATCTTAACATCGAAATGTTGCTATAGAC 5353  
QY 718 CAGTTGAAATCTCAGTCTCTCAGAGGAAAGCTGCTGACCAAGATTTCAAGGATTT 777  
DB 5354 CTGAGGCTTCTCAGATTTAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5413  
QY 778 GTGCTGTTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837  
DB 5414 GTGCTGATTTATTTCCAGGCACTTGAATACAGAAATCTATGACAGTATATTAAGTAG 5473  
QY 838 CAAGTGAAGCAATTAAGAACTGCAAAAGCTTCTGCGCTTAAGAGAGGCTTATTTGTG 897  
DB 5474 GAAATGATACAGGATGGAATATCTGCTGAGTGTGTAAGAGAGGATTTTATTTAGAG 5533  
QY 898 GAATATCTTCGAGAGCTGACGCTGCTGCTGCTTTCAGATTTGCAAAAAGACAGAAATG 957  
DB 5534 GTATTTATGAGGTGCTGCAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5593

RESULT 6

US-08-961-527-15  
Sequence 15, Application US/08961527  
Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunach  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10993 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-15

Query Match 18.4%; Score 250.2; DB 4; Length 10993;  
Best Local Similarity 57.6%; Pred. No. 3.2e-61;  
Matches 529; Conservative 0; Mismatches 378; Indels 12; Gaps 4;

QY 111 ATGCGAAGATGTTACGAGATGATGTTGTAACCCCATATGATATCTAAATAACT 170  
DB 7400 ATTATTAACAACATTAATGATTAATCGGTCAAGACCATGTTAACTTAACAACATC 7459  
QY 171 GCGGATGTTGTTGTTGCGCGGTTGCTGCTAAACTGAGTGTGATGAGCGATCTTACT 230  
DB 7460 GTCCGAGAGGTGCTGCAAGCGTATATTAAGCTTGAAGCATTATCTGTTCACT 7519  
QY 231 GTGAAGACAGATGTTGGTATAGTATGATGATGATGATGATGATGATGATGATGAT 290  
DB 7520 GTTAAAGACCTTATTCCTTCCCTTCAAGATGATGATGATGATGATGATGATGATGAT 7579  
QY 291 CCTGGAAGAGTGTCTCTCATTTAGCCACCAAGTGTGATATCTGCACTTTGATGATGAT 350  
DB 7580 CTTGTTTCTACT--ATTGTGAAGCAACAAGTGGAACCCGATTTGATGATGATGATGAT 7636  
QY 351 ATGGGACGAGCGAGGGGTTACAAGCTCATATTAACAAGCTGCTCTTATGATGATGAT 410  
DB 7637 GTAGGTGCTGCTAAAGGTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 7696  
QY 411 AGAAGATCATTTCTATTAGCTTTTGAAGCTGATGATGATGATGATGATGATGATGATGAT 470  
DB 7697 CGAGGTAATTAATTCACAGCTTATGTTGCTGAAGCTGCTCACTCTGTAAGCGAGGA 7756  
QY 471 ATGAAGGTGCTGTTTCAAGAGCTGAAGATTTTGGCTAAGCGCCCAATGCTTACATA 530  
DB 7757 ATGAAGGTGCTGTTTCAAGAGCTGAAGATTTTGGCTAAGCGCCCAATGCTTACATA 7813  
QY 531 CTTCAACAATTTGTAACCTGCGCAATCCCAAGTTCATTTATGAAACAATGCTGCTCAAG 590  
DB 7814 CTTCTTCATTTTGAATCAGTCAATTCAGAAAGTACAGAAAGAACACAGAGAGCTGAG 7873

QY 591 ATATGAAAGGCTCCG--ATGGAAAATTGATGATGATGATGATGATGATGATGATGATGAT 647  
DB 7874 ATATGATGCTGTTTCCGTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 7933  
QY 648 GGTCAATTAACAGGTGCTGGAATAATCTTAAAGACGAATCCGAATATAAGCTGAT 707  
DB 7934 GGAAGATTTTGTGTTTCTCATGACTCAATTCAGAAATTTTACATTCAGTTT 7993  
QY 708 GGTGGAACCAAGTTGAAAGTCCAGTCTTCAAGAGAAAGCTGTCCACACAGATT 767  
DB 7994 GCAGTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8053  
QY 768 CAAGGATGCTGCTGTTTATCCGAGTCTTGAAGCAATCTTGAAGATT 827  
DB 8054 CAAGGATGCTGCTGTTTATCCGAGTCTTGAAGCAATCTTGAAGATT 8113  
QY 828 GTTCAATATCAAGTGAAGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 887  
DB 8114 GTTGTATATCAAGTGAAGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 8173  
QY 888 CTAATTTGGAATATCTTCCGAGCTGCAAGCTGCTGCTTTCAGATGCAAAAGA 947  
DB 8174 TTCTTGTAGGATTTTCTCAGCTGCAAGCTGCTGCTTTCAGATGCAAAAGA 8233  
QY 948 CCAAGAAATGCGGGAAGCTTATGTTGCGGTTTCCAGCTTCCGAGAGAGTACTG 1007  
DB 8234 TTAG--GTACAGTAAAGATCTTTCCTTCCCTTACAGCAAGTAAGGTAAGTATCTC 8290  
QY 1008 TCTTCGCTGCTATTGACT 1026  
DB 8291 TCTACAGCACTTATGAT 8309

RESULT 7  
US-09-107-532A-271  
Sequence 271, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 710  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 893-5007  
TELEFAX: (781) 893-8277  
INFORMATION FOR SEQ ID NO: 271:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 927 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...927  
SEQUENCE DESCRIPTION: SEQ ID NO: 271;  
US-09-107-532a-271

Query Match 18.2%; Score 248.2; DB 4; Length 927;  
Best Local Similarity 58.0%; Pred. No. 3.1e-61;  
Matches 499; Conservative 0; Mismatches 353; Indels 9; Gaps 3;  
123 GTTACGGAATGATGTTGTAACCCCATGATATATCTAAATTAATCTGCGAGTGGT 182  
22 GTACCCGAGTTGATGCAAAACCCCAATGCTCAATTAATTAATGATACAGAGAT 81  
183 GTTCCCGGCTGCTGCTAACTGAGTGTAGAGCCAGCTCTAGTGTAGAGAGCAGG 242  
82 TCAGCAGATGTTTGTAAAGCTGAGTTTATATCAGAGAGAGAGTAAAGATGCT 141  
243 ATTGGATATAGTATGATGCTGATGCAAGAGAGAGAGAGCTTATCAACCTGGAAGAT 302  
142 ATTGCTTAAATGATGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 201  
303 GTCTCATGAGCCAGCAAGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362  
202 ---ATCATGAACTTCTCTGGAATACAGAGATGAGATGCTGATGAGAGAGAGTCA 258  
363 AGGGGTACAGAGCTCATATATTAATGATGCTGCTCTATGATCTTGAGAGAGAGAT 422  
259 AAGAGATATTAAGATATCATGCTGATGCTGAGAGAGAGAGAGAGAGAGAGAGAT 318  
423 CTATTAAGCTTTGAGAGCTGAGTGTGCTGAGAGAGATCTGCTTAAGAGAGAGAG 482  
319 ATGAAAGAGTACGCGAGAGATGATATTTGATCTGCTGAGAGAGAGATTTCTGATCA 378  
483 GTTCAAGAGCTGAGAGATGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAT 542  
379 AT---CAGAT 435  
543 GAAAG 602  
436 GAAATAG 495  
603 TCCGATGAG 659  
496 TTTGAG 555  
660 GGTGCTGAG 719  
556 GGCCTGAG 615  
720 GTTAAAG 779  
616 GCAAG 675  
780 GCTGATTTATCTCTGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839  
676 AAG 735  
840 AGTATGAG 899  
736 GGAAG 795  
900 ATATCTTCCGAG 959  
796 ATCTCATGAG 855

QY 960 GGAAGCTTATTTGAGCGCT 980  
DB 856 AAGAGATATGCTGTGCT 876

RESULT 8  
US-09-134-001C-493  
Sequence 493, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 493  
LENGTH: 960  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-493

Query Match 18.0%; Score 245.8; DB 4; Length 960;  
Best Local Similarity 57.6%; Pred. No. 1.5e-60;  
Matches 521; Conservative 0; Mismatches 372; Indels 12; Gaps 4;

QY 121 ATGTTACGAAATGATGTTGTAACCCCATGATATATCTAAATTAATCTGCGAGTGGT 180  
DB 50 ATGTTACGAAATGATGTTGTAACCCCATGATATATCTAAATTAATCTGCGAGTGGT 109  
QY 181 GTTCCCGGCTGCTGCTAACTGAGTGTAGAGCCAGCTCTAGTGTAGAGAGCA 240  
DB 110 ATGCAAGCTATATTTATGTTAGTATGATATCAAAATCAGAGTGGTGGTAAAGATC 169  
QY 241 GATTTGGTATATGATGATGCTGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
DB 170 GATATCTTAAAGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227  
QY 301 GGTGCTCATGAG 360  
DB 228 -TCAATCTTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 286  
QY 361 CCAAGGAGTACAGCTCATATTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
DB 287 CCAAGGAGTACAG 346  
QY 421 TTTATTAAGCTTTTGGAGTGTGAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 347 TATTAAGCTTTTGGAGTGTGAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406  
QY 481 CTGTTCAAG 540  
DB 407 CATTAAG 463  
QY 541 TTTAAAG 600  
DB 464 TCAAG 523  
QY 601 GCTCCAG 657  
DB 524 ATTTGAG 583  
QY 658 CAGTGTGAG 717  
DB 584 CTGTTGAG 643  
QY 718 CAGTTAAAG 777

Mon May 24 08:18:49 2004

us-09-931-457a-30.rml

Page 7

Db 644 CTGAGCTTCTCCGATTTAAGCGGTGTGAGACAGGCGCTCATTAATTAAGAGATTG 703  
Qy 778 GTGCTGTTTATCCCTGGGTCTTGGAAGTCAATCTTCTTAAGTGTCAATAT 837  
Db 704 GAGCAGTTTCTGACTGATCTTAAATACAGAGTTTATGACAGCATCATCAAGT 763  
Qy 838 CAAGTATGAGCAATAGAACTGCAAGCTTCTTCCGCTTAAGAAGCTATTTG 897  
Db 764 GTATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 823  
Qy 898 GATATCTTCCGAGCTGCTGCTGCTTCTTCAATGCAATGCAATGCAATG 957  
Db 824 GTATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881  
Qy 958 CCGGAGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017  
Db 882 -AGTAAACAGTTGTAAGATTTTCAAGTATGAGGAGTTCCTTATCAACCAT 940  
Qy 1018 TATTT 1022  
Db 941 TATAT 945

RESULT 9  
US-09-103-840A-2  
Sequence 2, Application US/09103840A  
Patent No. 6294328

GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, c, t, c or g  
US-09-103-840A-2

Query Match 16.5%; Score 225.2; DB 3; Length 4403765;  
Best Local Similarity 53.3%; Pred. No. 1.2e-52;  
Matches 498; Conservative 0; Mismatches 433; Indels 3; Gaps 1;

Qy 109 GAATTCGCAAGATGTTAGCAATTTGTTGTTAAACCCCATTAATTAATTAATTAAC 168  
Db 2604608 GCATCCCGAGGACATCACTCATCTGCGGCGCAACCGCTGCTGCGCGCGG 2604667  
Qy 169 TTGCGATGTTGTTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 228  
Db 2604668 TACCGACGCGCGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2604727  
Qy 229 GTGTGAGAGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 288  
Db 2604728 GGTAAAGACCTATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2604787  
Qy 289 CACCTGGAAGAGTGTCTCATTTAGCAACAGTGTATTTCTGCTGCTGCTGCTGCT 348  
Db 2604788 AGCCGACACGATCATTTCTC--GAACTCAAGCGCTGTAACAGCGCATCGCCCTG 2604844  
Qy 349 TCATGCGACGACGCGGCTTACAGACTTAATTAATTAATTAATTAATTAATTAAT 408  
Db 2604845 TGGTTGCGCGGACGCGGCTTACAGACTTAATTAATTAATTAATTAATTAATTAAT 2604904

Qy 409 AGAGAGATCATTTATTTAGCTTTTGAAGCTGAGTTGTTGCAAGATCTGCTAAG 468  
Db 2604905 AGCGCGAGATTTGCTGCGCGCATAGAGGTTGTAATCTCATCTGCTGCGGCTGCGAG 2604964  
Qy 469 GAATGAAAGTCTGTTTCAAGAGCTGGAAGATATTTGCTTAAGAGCCCATGCTTAA 528  
Db 2604965 GCATGTCAGTGCATTCGCAAGGCTGAGAGCTGCGCAAGCCCATCACTACTTCTG 2605024  
Qy 529 TACTTCAATTTTAAACCCCTGCAATCCCAAGTTTATTTGTAAGCACTGCTGCA 588  
Db 2605025 TGCCCGAGATTTGAGAGACCGGAGACCGGCGCATCATGCGCTGACGACCGGAG 2605084  
Qy 589 AGATGGAAGGCTCCGATGAGAAATTTGATGCTTTGTTTCTGAGATAGGCTGCTG 648  
Db 2605085 AGCTGCGGTGACACCGACGCAAGGTGCAATGCTGCTGCGGAGTCCGACCGG 2605144  
Qy 649 GTACATTAACAGTGTGGAATAATTTTAAAGACAGATTCGATATTAAGCTGATTG 708  
Db 2605145 GCATCATCACCGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605204  
Qy 709 GTGTGAAACAGTTGAAAGTCTCTCAAGAGAAAGCTGCTGCTGCTGCTGCTG 768  
Db 2605205 CCGTAGAGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605264

Qy 769 AAGGATTTGCTGCTGCTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828  
Db 2605265 AAGGATTTGCTGCTGCTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605324  
Qy 829 TTCAATTAACAGTGTGGAATAATTTTAAAGACAGATTCGATATTAAGCTGATTG 888  
Db 2605325 TTACGCTGCTGTAAGCAAGCGGCTCAAGTGTGCTGCTGCTGCTGCTGCTG 2605384  
Qy 889 TATTTGCGAATTTCTTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948  
Db 2605385 TCGTGTGCGATTTCTTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605444  
Qy 949 CAGAAATGCGGAGGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008  
Db 2605445 CAGAAATGCGGAGGCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605504  
Qy 1009 CTTGCTGCTATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1042  
Db 2605505 GCACCGCTGTTGCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605538

RESULT 10  
US-09-103-840A-1  
Sequence 1, Application US/09103840A  
Patent No. 6294328

GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 16.5%; Score 225.2; DB 3; Length 4411529;  
Best Local Similarity 53.3%; Pred. No. 1.2e-52;  
Matches 498; Conservative 0; Mismatches 433; Indels 3; Gaps 1;  
Qy 109 GAATTCGCAAGATGTTAGCAATTTGTTGTTAAACCCCATTAATTAATTAATTAAC 168

Mon May 24 08:18:49 2004

us-09-931-457a-30.mri

**ପୃଷ୍ଠା ୫**

Db	2608798	GCATCCGCCGAGACATGACCCAACTCATTCGGGCGCCACAACCGCTGCTCGAGTCCGCTGGCGCCAG	2608857
Qy	169	TTGGCGAATGGTTGTGTTGCTCCCGGGTGTGCTTAACCTGGAGTTAATGAGCCATGCTCTA	228
Db	2608858	TCACCCGACGGCGCGCTGTGCCACACATGCTGCGCAAGCTGGAAATTTTCAACCCGGGCAACA	2608917
Qy	229	GTGTGAAGGACAGGATTTGGGTATGATATGATTTGCTGATGCAAGAGAGAGAGGACTTATCA	288
Db	2608918	GCGTAAAGACCGATTCGGGGTTCGATGCTCCAAAGCGCGGACAGACAGGCTTGTATCA	2608977
Qy	289	CACCTGGAAAAGTGTCTCTCATTTAGCCCAAAATGGTAAATATCTGGATATGGATTACCT	348
Db	2608978	AGCCGGACACACATCATTTTC--GAACTCCAGACCGGTAAACACCGGACTTCGCTGACCA	2609034
Qy	349	TCAATGGACGACCGCAGGGGTACAAGCTCATAAATTAACATGCTCTGTATAGTCTTG	408
Db	2609035	TGGTTTGGCGGACGCGGGTACCGGATCGATGCTGACCATCCCGAGACGATGAGTCTGG	2609094
Qy	409	AGAGAGATCATTTCTATTAGCTTTTGGAGCTGAGTTGGTTTGAAGATCTCTGTAAG	468
Db	2609095	AGCGCGGATTTTCTTGCGCCATACGGATGCTGAATCTATCTCACTCCGCGGTGGAAG	2609154
Qy	469	GAATGAAGGCTGCTTTCAAGAAGCTGAGAGATATTTGGCTAAGACGCCATGCTTACA	528
Db	2609155	GCAATGACGATGCCATTCGCAAGGCTGAGAGCTGGCCAGACCGATCAACGCTACTTGG	2609214
Qy	529	TACTTCAACATTTGAAAACTCTGCGCATTCCCAGGTTCAATTAGAAACACTGGTCCAG	588
Db	2609215	TGCCCCACCAATTCAGAAACCCCGGGAACCCGGCATCATCGGTCACGACCCGCGAG	2609274
Qy	589	AGATATGAAAAGGCTCCGATGGGAAATATGATATGCAATTGTTCTGGATAGGCACTGGT	648
Db	2609275	AGGTGTGGCGTAAACCGACCGCAAGTGCAGCATGTCGTGCGCGGAATCGGACCCGGTG	2609334
Qy	649	GTAACAATACAGGTCTGGAAAAATCTTAAAGACAGAAATCCGAATATTAAGCTGATTG	708
Db	2609335	GCACCATACCGGCGCTCGCGCAGGTCATCAAGGAACGAAACCGTCGGCCCGGTTCTGG	2609394
Qy	709	GTGTGAACACAGTTGAAAAGTCAGATGCTCTCAGAGAGAAAGCTGGTCAACAAGATTC	768
Db	2609395	CGGTAGACCCCGGCGCGTCCGCGGCTCTTTCTGGTGACCAAGGGAACCGCACCCGATCC	2609454
Qy	769	AAGGAAATGGTGTGCTTTTATCTCTGATGCTTGGAAAGTCAATCTTGTATAGAAGTTG	828
Db	2609455	AAGGATATCGGCGCCGAGTTCTGCTCCGCGGATCTGACACAGAACTTAATCGACGATCA	2609514
Qy	829	TTCAATATTCAGTATGATGAGCAATATAGAAACTGCAAGCTTCTTGGCTTAAAGAGGCC	888
Db	2609515	TTACCGTGGTAAACGAAGCGGCTCAACGTGGCGCGCGAGTGGCCCGGAAAGAGGGCT	2609574
Qy	889	TATTTGTGGGAATATCTTCCGAAGTGCAGCTGCTGCTTTCATGATCCAAAAAAC	948
Db	2609575	TGCTGTGCGGCAATCTCTCGGGGCGCGCCACAGTGGCCGCTCTTAAGGTGGCCGCGCGC	2609635
Qy	949	CAGAAAATGCCCCGGAAGCTATTTGTGCGGTTTTCAGCTTCGGGAGAGAGTACTGT	1008
Db	2609635	CAGACAAACGCCGGAAGCTAATCGTGTATGTCTCCCGACTTCGGCGAACAAGATATCTGA	2609694
Qy	1009	CCTCCGTGCTATTTAGTCAGTGAACCGGAAGC	1042
Db	2609695	GCACACCGTTGTGCGCGACGTGGCTGACTAAGC	2609728

RESULT 11  
 US-09-557-884-1  
 Sequence 1, Application US/09557884  
 Patent No. 6506581  
 GENERAL INFORMATION:  
 APPLICANT: Fleischmann et al.  
 TITLE OF INVENTION: The Nucleotide sequence of  
 the Haemophilus influenzae Rd Genome, Fragments  
 thereof, and Uses Thereof

```

1 NUMBER OF SEQUENCES: 1
2 CORRESPONDENCE ADDRESS:
3 ADDRESSER: Human Genome Sciences, Inc.
4 STREET: 9410 Key West Avenue
5 CITY: Rockville
6 STATE: MD
7 COUNTRY: USA
8 ZIP: 20850
9
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: 3 1/2 inch diskette
12 COMPUTER: Dell Pentium
13 OPERATING SYSTEM: MS DOS V6.22
14 SOFTWARE: Ascii Text
15
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/09/557,884
18 FILING DATE: 25-Apr-2000
19 CLASSIFICATION: <Unknown>
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 08/476,102
23 FILING DATE: JUN-5-1995
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Michelle S. Marks
26 REGISTRATION NUMBER: 41,971
27 REFERENCE/DOCKET NUMBER: FB186P3
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: 301-309-8504
30 TELEFAX: 301-309-8439
31
32 INFORMATION FOR SEQ ID NO: 1:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 1830121 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37 TOPOLOGY: linear
38
39 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
40 US-09-557-884-1
41
42 Query Match 16.4%; Score 223.8; DB 4; Length 1830121;
43 Best Local Similarity 57.5%; Pred. No. 1,9e-52;
44 Matches 493; Conservative 0; Mismatches 337; Indels 27; Gaps 4;
45
46 192 GTTGCTCTTAAGCTGAGTGTGATGAGACCCATGCTCTTAGTGTGAAGACAGATTGGGTAT 251
47 1165211 GTGATGTGTAAATAATGAAAGCTGTGATACCCAGCTACAGCGTAAATGCGATTGGGCGC 1165270
48
49 252 AGATGATTTGCTGTGACAGAGAGAGAGAGACTTATCACCTGGAAAGAGTGTCTTATT 311
50 1165271 AATATGATGTGGCAGACGAGAAAAAGATGCAACCTCCACAAAAG--GAAAGAGATTGTA 1165327
51
52 312 GAGCCAAACAGTGTAACTACTGCGATTGATTAACCTTCATGTCAGACAGCCAGGGATTAC 371
53 1165328 GATCAACAGATGTGTAACACGGGATTCCTTTGCTTATGTTGGGGGCTCAAGGTTAT 1165387
54
55 372 AAGCTCAATTAATCAAGCCGCTCTTATAGCTTGTGAGAGAAATCATTTATTAGCT 431
56 1165388 AAATATCGCTCACTATGCCAGAAACATATAGCTTTGAAAGAAAAGCTTATGTGGGA 1165447
57
58 432 TTGAGACTGAGTTGGTTTCTGACAGATCTGCTTAAGGAAATGAAGGTGCTTTCAGAG 491
59 1165448 TTGGGTGTAAATTATAGCTTACCGAAGGGCGCAAAAGAAATGAAGGTGCTATTTGCGAAA 1165507
60
61 492 GCTGAAGAAGATATTGGCTTAAGACGCCCAATGCTTA--CATACTTCAACATTTGAAAC 548
62 1165508 GCGAAGAAATTTGTGCTTCTGATCCAAAGCGGTATGTATGCTTAACAATTTGAAAT 1165567
63
64 549 CTTGCCAATCCCAAGTTCAATTATGAACAATCTGTCTCAAGAGATATGGAAGAGCTCCGAT 608
65 1165568 CCAAGCCAAACCAAAATTCATCGAGAAACAAGGGCTTAATTTGAAAAGATACAGAT 1165627
66
67 609 GGGAAATATGATGATTTGTTTCTGGATATAGGCACTGTGTGATCAATTAACAG--TGCT 665
68 1165628 GCGAAATTCAGATGTGTGTTGCTGTGGGTAGGAACAGTGTGTTGATTAACGGGCAATTCT 1165687
69
70 666 GGAATATATCTTAAGACCAATCCGAATATTAAGCTGATGTGTGTGGAACAAGTTGAA 725

```



Sequence 1, Application US/09790988  
Patent No. 6632935  
GENERAL INFORMATION:  
APPLICANT: SHIGENOBU, SHUJI  
APPLICANT: MATANABE, HIDEMI  
APPLICANT: HATTORI, MASAHIRA  
APPLICANT: SAKAKI, YOSHIYUKI  
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
FILE REFERENCE: 081356/0159  
CURRENT APPLICATION NUMBER: US/09/790,988  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: JP2090-107160  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1  
LENGTH: 640681  
TYPE: DNA  
ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 16.1%; Score 218.8; DB 4; Length 640681;

Best Local Similarity 56.2%; Pred. No. 2.9e-51;

Matches 534; Conservative 0; Mismatches 377; Indels 39; Gaps 5;

Db 100 AAGGTCGGAATGCGCAAGATGTTACGAATTGATTGTGTAACCCCATAGTATATC 159  
72179 AATGAGTAAATATATCAAGATATTTCTTACATTTGTAATACCTTACTGCTT 72120  
Qy 160 TAAATTAACCTGCGATGTTGTGTCGCGGTTGCTGCTAACTGAGTTGAGAC 219  
72119 TAAATGAGATGCAATGGAAT-----ATTGATGTAATATGAACTTGAATC 72069  
Qy 220 CAGCTCAGAGTGAAGACAGAGTTGGTATAGATGATGCTGAGCGAAGAGAGG 279  
72068 CAGCTTATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 72009  
Qy 280 GACTTATCACACCTGGAAGAGTGTCTCATGAGCCCAAGTGTATATCTGGCATG 339  
72008 AATATATATA---TAAATATGTAATGTAATGTAATGTAATGTAATGTAATG 71952  
Qy 340 GATTAGCTTATGAGCAGACGAGGAGGTTACAGCTCATATATGCAATGCTCTCTA 399  
71951 CATTAGCTTATGAGCAGCTTCTGAAATATGATTAATCTTAACTGATGCTGAAAC 71892  
Qy 400 TGAGCTTGAAGAGATCATCTTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 459  
71891 TGTCTATGAGCGGAAATATTAATTAATCTTTAGTGAGCAATGTTATTTACCATG 71832  
Qy 460 CTGCTAAGGATGAAAGTGTGCTTCAAGAGCTGAGAGATATGCTTAAAGCCCA 519  
71831 GAAATATGAGATGAAAGGAGCTATTTCTAAGGCAATGATATATCTGCAATCCA 71772  
Qy 520 ATGCTA---CATCTTCAACAAATTTGAAACCTGCAATCCCAAGTTCAATATGAA 576  
71771 GTAAATATTTTATTAACAATTTGAAATCCAGCTATCTGAAATTCATCAATTA 71712  
Qy 577 CCACTGTCCAGAGATGAGAAAGGCTCCAGTGGGAAATGATGCAATTTTCTGGA 636  
71711 CTACTGAGCAGAAATATGAGACGATATGCAATGCAATTTGATTAATTTCTGCTG 71652  
Qy 637 TAGGACCTGAGTGAATCAATCAAGTGTGAGAAATATCTTAAGAGCAGATCCGATA 696  
71651 TAGGACCGAGTGCACATATCTGATATCAAGATATATCAAAAAATTAAGAAAAA 71592  
Qy 697 TAAAGCTATTTGCTGTG-----GAACTGTAAGATGCTGCTCTGAG 741  
71591 AAAATTTATAGATATAGCTGTGGAACCTTCTGAATCACTGATATCAACAATTTTATG 71532  
Qy 742 GAGGAA-----AGCTGTCCAGCAGATTCAGAGATGAGGATGAGTGTGTTATCC 792  
71531 CAGGAAAGCATGAGCTGTGACCAATTAATTTCAAGAAATTTGGGCTGTGTTATAC 71472

Qy 793 CTGCTGCTTGAAGCATCTTCTGATGAAGTGTGCAAAATATCAAGTGAAGCA 852  
Db 71471 CCAAAATCTAGATTTTAACAAATATGATCAAGTATTAATCTTTCTAGGAAAGCA 71412  
Qy 853 TAGAACTGCAAGCTTCTTGGCTTAAAGAGCCATTTTGGGAATATCTTCCGAG 912  
Db 71411 TATTAAGCTTAAGATTAATGAAGAAAGAGTATCTAGCAGAAATCTTCTGCTG 71352  
Qy 913 CTGCAAGCTGCTGCTTCTTCAATTCGAAAGAAACAGAAATGCGGGAAGCTATTG 972  
Db 71351 CTGCTTATATGCAAGATTAATTAATCAACAGCAAAAAATTTTCAAGCAAAAAATAG 71292  
Qy 973 TTGCGCTTTTCCAGCTTCCGAGAGAGGATACCTGCTCCGCTATT 1022  
Db 71291 TTGTTATTTTACCTTCTTCAAGGAAAGCTTTATTAATGATCAAGATTAAT 71242

## RESULT 14

US-09-724-623-10

Sequence 10, Application US/09724623

Patent No. 6476209

GENERAL INFORMATION:

APPLICANT: Glenn, Matthew

APPLICANT: Lubbbers, Mark W

APPLICANT: Dekker, James

TITLE OF INVENTION: Polynucleotides, materials incorporating

TITLE OF INVENTION: them, and methods for using them.

FILE REFERENCE: 1048U1

CURRENT APPLICATION NUMBER: US/09/724,623

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 124

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10

LENGTH: 927

TYPE: DNA

ORGANISM: Lactobacillus rhamnosus

US-09-724-623-10

Query Match 15.4%; Score 209.8; DB 4; Length 927;

Best Local Similarity 55.3%; Pred. No. 3.1e-50;

Matches 472; Conservative 0; Mismatches 372; Indels 9; Gaps 3;

Qy 114 GCCAAGATGTTACGATGATGTTGTAACCCCAATTTGATATCTTAATAACTGCG 173  
Db 13 GCGATATATTTCAAGTTTATTTGCAATGCGCCCTACTCAAGCTCAATCCGCTTGT 72  
Qy 174 GATGTTGTTGTCGCGGTTGCTGTAATCTGAGTTGAGGCAATGCTCTAGTGT 223  
Db 73 CTGAAGGCGCGCGATGTTTATGTCAGCTGAAATCTTTAATCCGCGGCTCAGTC 132  
Qy 234 AAGACAGATTTGGATATGATGATGCTGATGCAAGAGAGAGGAGCTTATCACACT 293  
Db 133 AAGACCGAGATGCTTGGGATGATGAAAGCGCTGATATTAAGGGGCTCTGAAGCA 192  
Qy 294 GGAAGAGTGTCTCATTTGAGCAACAAGTGTATATCTGCAATGATGATGATGCTTCAG 353  
Db 193 GG---CGGCACTATGTTGAGCAAGTCCGCAACCGGCAATGCAATGCAATGCTGAT 249  
Qy 354 GCAAGACCGAGGTTTCAAGCTCATATTAACAATGCTCTTCTATGAGCTTGAAGA 413  
Db 250 GCGGCGCAAAAGTTATCACTCATATCAATGCGGAAACAGATGATGATGAGCGG 309  
Qy 414 AGAATCATTTTATGATTTTGAAGTGAAGTGTGCTGACAGATCCTGCTAAGGAGATG 473  
Db 310 CGGCTTTATATGATGTTATGAGAGCGAATCATTTTGAAGCGCGGATGCGATGAGATG 369  
Qy 474 AAGGCTGTTTGAAGGTTGAAGATTTTGGCTTAAGAGCGCAATGCTTCAATCTT 533  
Db 370 CCGGAGCAATTTAAAGAGAGAG---CATTAAGCAAGAAATGAGCTTACTTGGCA 426  
Qy 534 CAACAATTTGAAGACCTGCAATCCCAAGTTCAATTAAGAAACAATGCTGCTCAGA---G 590  
Db 427 ATGCATTTCAAGAACCCCGCAATTCAGAGTGTCAAGAGGCAAGCCGCAAGAAATC 486

QY 591 ATATGAAAGGCTCCGATGAGGAAATGATGATTTGTTTGGGATAGGCACTGTGTGT 650  
 DB 487 ATCCGTTCATTTGATGATGAGGCAACCCAGATGCTTTGTTAGCCGGGCGGCAAGGCGGA 546  
 QY 651 ACAATTAACAGTGTCTGAGAAATATCTTAAGAGCAGATCCGATATTAAGCTGATGTGT 710  
 DB 547 ACACATCACCGGGGTGTGTGGGCTCTGCGTAAGATCAATCCAGATGTACAAATCTATGTGG 606  
 QY 711 GGGAAACAGTTGAAGTCCAGTCTCAGAGGAAAGCTGTCTCACAAGATTTCA 770  
 DB 607 TTGGAGACGCGGATGCGCATGCTTAAGAAAGGCAATGGGCGGCAAGACAAATTTCA 666  
 QY 771 GGGATGTGTGTGTTTATCCCTGTGTCTTGGAAATCAATCTTCTTATGATTAAGTTGT 830  
 DB 667 GGGATCTCAGCCGGTTTATTCAGAGCTTATGATACGACCTTATCAGACATCAAT 726  
 QY 831 CAATATCAAGTATGATAGCAATAGAACTGCAAGCTTCTGGGCTTAAGAGGCGCTA 890  
 DB 727 GAAGTACACGAGATCAAGCTATGACATGCTCGCCAGTCAAGCCATAGAGAGGCTTC 786  
 QY 891 TTGTGGGAAATATCTTCCGAGCTGAGCTGTGTCTTTTCAATTGCAAAAGACCA 950  
 DB 787 CTACCGGAGATTTCCGCTGTGGCTTACATTTTGGCGGATTTGAATGCAAGAAATCTC 846  
 QY 951 GAAATGCCGCGGA 963  
 DB 847 GGCAGAGGCAAGA 859

RESULT 15

US-09-634-238-26/c  
 : Sequence 26, Application US/09634238  
 : Patent No. 6544772  
 : GENERAL INFORMATION:  
 : APPLICANT: Glenn, Matthew  
 : APPLICANT: Havukkala, Ilkka J.  
 : APPLICANT: Bloksberg, Leonard, N.  
 : APPLICANT: Eubbers, Mark W.  
 : APPLICANT: Dekker, James  
 : APPLICANT: Christenson, Anna C.  
 : APPLICANT: Holland, Ross  
 : APPLICANT: O'Toole, Paul W.  
 : APPLICANT: Reid, Julian R.  
 : APPLICANT: Coolbear, Timothy  
 : TITLE OF INVENTION: Polynucleotides, materials incorporating  
 : TITLE OF INVENTION: chem and methods for using them.  
 : FILE REFERENCE: 11000.1043U1  
 : CURRENT APPLICATION NUMBER: US/09/634,238  
 : CURRENT FILING DATE: 2000-08-08  
 : NUMBER OF SEQ ID NOS: 422  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 26  
 : LENGTH: 8885  
 : TYPE: DNA  
 : ORGANISM: Lactobacillus rhamnosus  
 US-09-634-238-26

Query Match 15.4%; Score 209.8; DB 4; Length 8885;  
 Best Local Similarity 55.3%; Pired. No. 1.1e-49;  
 Matches 472; Conservative 0; Mismatches 372; Indels 9; Gaps 3;

QY 114 GCCAAGATGTTACGATGATGTTGTTAAACCCCATATGATATCTTAATAAATCTTGGC 173  
 DB 5600 GCGATATATTTACAGTTTAAATGGCAATAGCCCGCTACTCAAGCTCAATCGCTTGTGA 5541  
 QY 174 GATGTTGTGTGTCGCGGCTGTCTTAACCTGATGATGAGCGCATGCTTATGTGTG 233  
 DB 5540 CCGAAGGCGCGCGGATGTTATGTCACAGCTGAATTTCTTAATCCGCGGCTCAGTC 5481  
 QY 234 AAGGACAGATTTGGATATAGATGATTTGATGACAGAAAGAGAGGACTTATCAGACCT 293  
 DB 5480 AAGGACAGATTTGCTTGGGATGATTTGAAGAGCGCTGAATATTAAGGCGCTTGAAGCA 5421

QY 294 GGAAGAGTGTCTCTCATTTAGCCCAAGAGTGTAAATATCTGCAATTTGATTAACCTTCTG 353  
 DB 5420 GG---CGGACACATTTGTTGAGACGTCGCGCAACCGGCATTTGAGCTGCACTGTGTT 5364  
 QY 354 GCGAGGCAAGGCGTTTCAAGCTCATATTTCAATGAGCTGCTTCTATAGATCTTATAGAGA 413  
 DB 5363 GCGGCGGAAAGGTTATACCTTATCATTCACATTCGCGGAAACGATGATGTTGAGCGG 5304  
 QY 414 AGAATCATTTATTAAGCTTTTGGAGCTGAGTGTGTTCTGACAGATCTGCTTAAGGAAATG 473  
 DB 5303 CGTGTGTTGATGCGGTTTACGAGCGGACATCAATTTGAGCGCGGATGCGATGAAATG 5244  
 QY 474 AAGGTGTCTTCAAGAGGCTGAGAGATTTGGCTTAAGAGCCCAATGCTTACTACT 533  
 DB 5243 CCGGAGCAATTTAAAGAGCAAGG---CATTAAGCAAGAAATGCTTACTTCTTGCA 5187  
 QY 534 CAACATTTGAAACCTGCAATCCCAAGTTCATTAAGAAACCACTGTGTCAGAG---G 590  
 DB 5186 ATGCAATTCAGAAACCCCGCATTCAGAGCTCCAGAGGCGGAGAGCCGACAAAGAAATC 5127  
 QY 591 ATATGAAAGGCTCCGATGAGGAAATGATGATTTGTTTCTGGAATGCACTGTGTGT 650  
 DB 5126 ATCCGTTCATTTGATGATGAGGCAACCCAGATGCTTTGATGAGCGGCGGCGCAAGGCGGA 5067  
 QY 651 ACAATTAACAGTGTGAGAAATATCTTAAGAGCAATCCGATATTAAGCTGATTTGT 710  
 DB 5066 ACATCTACCGGGGTGTGTGGGCTGTGCTTAAGATCAATTCAGATGATCAATATCATGCG 5007  
 QY 711 GTGAAACAGTGTGAAGTCCAGTGTCTCAGAGAGGAAAGCTGTCTCACAAGATTTCA 770  
 DB 5006 TTGGAAGAGCGGATGCGCATGCTTAAGAAAGCCATGCGGCAAGACCAAGATTTCA 4947  
 QY 771 GGGATGTGTGTGTTTATCCCTGTGTCTTGGAAATCAATCTTCTTATGATTAAGTTGT 830  
 DB 4946 GGGATCTCAGCCGGTTTATTCAGAGCTTATGAGAGAACTCTTATCAAGATCATTT 4887  
 QY 831 CAATATCAAGTATGAGCAATAGAACTGCAAGCTTCTTCCGCTTAAGAGGCGCTA 890  
 DB 4886 GAAATCACAGGATCAAGCTATGACATGCTGCGCACAGTCAAGCATGAAAGAGCTTC 4827  
 QY 891 TTGTGGGAAATATCTTCCGAGCTGAGCTGTGTCTTTCAATGCGAAAGAGCA 950  
 DB 4826 CTACCGGAGATTTCCGCTGCGCTTAACATTTTGGCGGATTTGAATGCGCAAGAAATCTC 4767  
 QY 951 GAAATGCCGCGGA 963  
 DB 4766 GGCAGAGGCAAGA 4754

Search completed: May 21, 2004, 20:35:46  
 Job time : 159 secs



Mon May 24 08:18:49 2004

us-09-931-457a-30.rnpb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2004, 19:00:02 ; Search time 703 Seconds  
(without alignments)

8805.006 Million cell updates/sec

Title: US-09-931-457a-30

Perfect score: 1362  
Sequence: 1 actctgtatgcctgtagatag.....aaaaaaaaaaaaaaaa 1362

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US04\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US03\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US02\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US01\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US00\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US00A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US00B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US00C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US00D\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US00E\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US00F\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US00G\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US00H\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US00I\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US00J\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US00K\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	100.0	1362	US-09-931-457a-30	Sequence 30, Appl
2	1340.6	98.4	1407	US-10-424-559-130460	Sequence 110460, A
3	660.8	48.5	3075	US-10-424-559-55164	Sequence 55164, A
4	595.2	42.7	966	US-09-938-842a-893	Sequence 893, App
5	595.2	42.7	966	US-09-938-842a-893	Sequence 893, App
6	578.6	42.5	1404	US-10-424-559-55160	Sequence 55160, A
7	578.6	42.5	1404	US-10-424-559-55160	Sequence 55160, A
8	557.4	40.9	1319	US-10-425-114-15467	Sequence 15467, A
9	555.8	40.8	1312	US-10-425-114-15467	Sequence 15467, A
10	555.8	40.8	1312	US-10-425-114-15467	Sequence 15467, A
11	543.4	39.9	1603	US-10-424-559-61234	Sequence 61234, A
12	543.4	39.9	1603	US-10-424-559-61234	Sequence 61234, A
13	524.4	38.5	1490	US-10-425-114-22525	Sequence 22525, A
14	521.8	38.3	1490	US-10-425-114-22525	Sequence 22525, A

15	521.8	38.3	1491	US-10-425-114-34527	Sequence 34527, A
16	521.6	38.3	1522	US-09-931-457a-63	Sequence 63, Appl
17	520.2	38.2	1508	US-09-931-457a-61	Sequence 61, Appl
18	520.2	38.2	1509	US-10-425-114-30493	Sequence 30493, A
19	495.8	36.4	975	US-09-905-290A-3	Sequence 3, Appl1
20	477.2	35.0	972	US-09-938-842a-495	Sequence 495, App
21	477.2	35.0	972	US-09-938-842a-495	Sequence 495, App
22	477.2	35.0	1472	US-09-887-576-611	Sequence 611, Appl
23	476.4	35.0	1329	US-09-938-842a-99	Sequence 99, Appl
24	476.4	35.0	1329	US-09-938-842a-99	Sequence 99, Appl
25	474.8	31.9	2307	US-10-424-559-53377	Sequence 18037, A
26	474.8	31.4	1925	US-10-424-559-53377	Sequence 53377, A
27	396.8	28.7	594	US-10-424-559-47440	Sequence 47440, A
28	396.8	28.7	594	US-10-424-559-47440	Sequence 47440, A
29	385.4	26.0	808	US-10-424-559-18732	Sequence 18732, A
30	353.6	24.6	528	US-09-770-152-879	Sequence 879, App
31	335.2	24.2	1095	US-10-424-559-99835	Sequence 99835, A
32	329.6	24.2	930	US-10-282-122a-12236	Sequence 15848, A
33	327.2	24.0	948	US-10-424-559-103489	Sequence 103489, A
34	322.6	23.7	962	US-09-770-445-305	Sequence 305, App
35	308	22.6	927	US-10-282-122a-16019	Sequence 16019, A
36	303.2	22.3	927	US-10-282-122a-16019	Sequence 652, App
37	300.4	22.1	542	US-09-770-152-652	Sequence 109938, A
38	294	21.6	1021	US-10-424-559-109938	Sequence 12236, A
39	280.8	20.6	945	US-09-734-017a-27	Sequence 27, Appl
40	276.2	20.3	568	US-10-282-122a-26908	Sequence 26908, A
41	273	20.0	927	US-10-672-787-39	Sequence 39, Appl
42	273	20.0	100848	US-10-424-559-55162	Sequence 55162, A
43	262.6	19.3	529	US-10-424-559-127232	Sequence 127232, A
44	256.6	19.0	373	US-10-282-122a-21035	Sequence 21035, A
45	256.4	18.8	930	US-10-282-122a-21035	Sequence 21035, A

# ALIGNMENTS

RESULT 1  
US-09-931-457a-30  
Sequence 30, Application US/09931457A  
Parent No. US20020157121A1  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
FILE REFERENCE: B1116 US CIP  
CURRENT APPLICATION NUMBER: US/09/931,457A  
CURRENT FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: 09/424,976  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: 60/065,385  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/049,406  
PRIOR FILING DATE: 1997-06-12  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 30  
LENGTH: 1362  
TYPE: DNA  
ORGANISM: Glycine max  
US-09-931-457a-30

Query Match  
Best Local Similarity 100.0%; Score 1362; DB 9; Length 1362;  
Matches 1362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTGTAGTTCGAGATAGCCATGCTGTCTTATGATGTCAGTCACTTCTTCT 60  
DB 1 ACTTGTAGTTCGAGATAGCCATGCTGTCTTATGATGTCAGTCACTTCT 60  
QY 61 CAAGTCAAGCTTGTAGTACGATATATATGCTGTGAAAGTCCGGAATGCCAAG 120  
DB 61 CAAGTCAAGCTTGTAGTACGATATATATGCTGTGAAAGTCCGGAATGCCAAG 120

QY	121	ATGTTACGGAAATGATTTGTAATAAACCCCAATAGTGTATCTAAATTAATCTGACGGATGGTT	180
Db	121	AGTTTACGGAAATGATTTGTAATAAACCCCAATAGTGTATCTAAATTAATCTGACGGATGGTT	180
QY	181	GTGTGCCCCGAGTGTGCTGTAAACTGGAGTTGATGAGCCCATGCTCTAGTGTGAAGCA	240
Db	181	GTGTGTGCCCGGAGTGTGCTGTAAACTGGAGTTGATGAGCCCATGCTCTAGTGTGAAGCA	240
QY	241	GGAATTGGGTATATGTATGATGCTGTATGCGAAGAAAGAAAGGACTTATCAACTCTGAAAGA	300
Db	241	GGAATTGGGTATATGTATGATGCTGTATGCGAAGAAAGAAAGGACTTATCAACTCTGAAAGA	300
QY	301	GTGTCCCTAATGAGCCAAACAAGTGTATATCTGCGATTGATGATTTAGCCCTTCACTGGCAGCAG	360
Db	301	GTGTCTCTAATGAGCCAAACAAGTGTATATCTGCGATTGATGATTTAGCCCTTCACTGGCAGCAG	360
QY	361	CCAGGGGTTTACAAGCTCATTAATTACAATGCTGCTCTTATGATGCTTTAGAGAAAGATCA	420
Db	361	CCAGGGGTTTACAAGCTCATTAATTACAATGCTGCTCTTATGATGCTTTAGAGAAAGATCA	420
QY	421	TTCTATTATGCTTTTGGAGCTGAGTGGTGTCTTGACAGATCCTGCTTAAGGAATGAAGGTG	480
Db	421	TTCTATTATGCTTTTGGAGCTGAGTGGTGTCTTGACAGATCCTGCTTAAGGAATGAAGGTG	480
QY	481	CTGTTCAGAAAGCTTGAAAGATATTTGGCTTAAGAAGCCCAATGCTCTACATCTTCAACAT	540
Db	481	CTGTTCAGAAAGCTTGAAAGATATTTGGCTTAAGAAGCCCAATGCTCTACATCTTCAACAT	540
QY	541	TTGAAAACCCCTGCCAATCCCAAGCTTCAATPNTGAAACAACGTGTCCAAAGATATGGAAG	600
Db	541	TTGAAAACCCCTGCCAATCCCAAGCTTCAATPNTGAAACAACGTGTCCAAAGATATGGAAG	600
QY	601	GCTCCGATGGGAAATATGATGATTTGTTCTGGGATAGGCACTGTGTGTATCAATTAACAG	660
Db	601	GCTCCGATGGGAAATATGATGATTTGTTCTGGGATAGGCACTGTGTGTATCAATTAACAG	660
QY	661	GTGCTGGAAAATATCTTAAAGACAGAAATCCGAATATTAAGTGAATTTGTGTGAACAG	720
Db	661	GTGCTGGAAAATATCTTAAAGACAGAAATCCGAATATTAAGTGAATTTGTGTGAACAG	720
QY	721	TTGAAAGTCCAGTGTCTCAGAGGAAAGCCCTGGTCCACAACAAGTTCAAGGCAATTTGGTG	780
Db	721	TTGAAAGTCCAGTGTCTCAGAGGAAAGCCCTGGTCCACAACAAGTTCAAGGCAATTTGGTG	780
QY	781	CTGGTTTATCCCTGTGTCTTGGAAGTCAATCTTCTTGAAGAGTTGTTCAATATCA	840
Db	781	CTGGTTTATCCCTGTGTCTTGGAAGTCAATCTTCTTGAAGAGTTGTTCAATATCA	840
QY	841	GTGATGAGCAATAGAACTGCAAGCTTCTGTGACCTTAAGAAAGGCTATTTGTGGAA	900
Db	841	GTGATGAGCAATAGAACTGCAAGCTTCTGTGACCTTAAGAAAGGCTATTTGTGGAA	900
QY	901	TATCTTCGGAAGCTGACGTGCTGTCTTTTCAGATTGCAAAAAGCCAGAAATATCCG	960
Db	901	TATCTTCGGAAGCTGACGTGCTGTCTTTTCAGATTGCAAAAAGCCAGAAATATCCG	960
QY	961	GGAAGCTTAATTTGTGCGGTTTTTCCACGCTTGGGGAAGAGTACTGTCTCCGTGTAT	1020
Db	961	GGAAGCTTAATTTGTGCGGTTTTTCCACGCTTGGGGAAGAGTACTGTCTCCGTGTAT	1020
QY	1021	TTGATCTACGTAGAGCGGAGCTGAAGCATGATTTGAGCCTCTGAATTCGCGTTTAA	1080
Db	1021	TTGATCTACGTAGAGCGGAGCTGAAGCATGATTTGAGCCTCTGAATTCGCGTTTAA	1080
QY	1081	GCTCTCACTACGTAAATTTTCTGTTAATTTGTAACGAGCTTTTATACGATTTCTTGAAGAC	1140
Db	1081	GCTCTCACTACGTAAATTTTCTGTTAATTTGTAACGAGCTTTTATACGATTTCTTGAAGAC	1140
QY	1141	TACTGTTTGTGACTGTGACTCTAAATTAATTAATCTTGCTCCAAAGAATAAGTTTCTTGAT	1200
Db	1141	TACTGTTTGTGACTGTGACTCTAAATTAATTAATCTTGCTCCAAAGAATAAGTTTCTTGAT	1200
QY	1201	GCCCTGAGAGCATATTTTGTGCTGCAACATTTAAAGATATTCAAAGTTGCTTATPAG	1260

Db	1201	GCCCTTGGAGCGATTAATTTTGTGCTGCACATTTAAATAAGTATCAAGTTCCTATTAAG	1260
Cy	1261	TAACATGTTTCATCTTTTGTGTGTTGAGACGACACGAGTGAAGTCAATATCTATGT	1320
Db	1261	TAACTGTTTCATCTTTGTGTGTTGAGACGACACGAGTGAAGTCAATCTATGT	1320
Cy	1321	TTTCGATTTCTTGTGTAGGAAAAAAAAAAAAAAAAAAAA	1362
Db	1321	TTTCGATTTCTTGTGTAGGAAAAAAAAAAAAAAAAAAAA	1362

```

RESULT 2
US-10-424-599-130460
: Sequence 130460, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: Le Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 130460
: LENGTH: 1407
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURES:
: OTHER INFORMATION: Clone ID: PAT_MRT3847-88813C.1
US-10-424-599-130460

```

Query Match	98.4%;	Score 1340.6;	DB 13;	Length 1407;
Best Local Similarity	99.3%;	Pred. No. of		
Matches 1346;	Conservative	0;	Mismatches	9;
			Indels	0;
			Gaps	0;
Qy	2	CTTGTGAGTGGATAGACCGATGTGCTGTCTTAATGTGTCAAGTCATTCGTGCTC	61	
Dp	38	CTTTGTAGTGGATAGACCGATGTGCTGTCTTAATGTGTCAAGTCATTCGTGCTC	97	
Qy	62	AAATCAAGCTTTGTAGTGAACAGATATTAATGCTTTGAAGAAGTCCGAAATTCGCAAGA	121	
Dp	98	AAATCAAGCTTTGTAGTGAACAGATATTAATGCTTTGAAGAAGTCCGAAATTCGCAAGA	157	
Qy	122	TGTTACGAAATGATTTGTTAAACCCCATTAAGTATCTAAATAACTTCGATGTTG	181	
Dp	158	TGTTACGAAATGATTTGTTAAACCCCATTAAGTATCTAAATAACTTCGATGTTG	217	
Qy	182	TGTTGCCCCGGTGTCTCTAACTGAAGTTGAATGAGCCATGCTCTAGTGTGAAGACAG	241	
Dp	218	TGTTGCCCCGGTGTCTCTAACTGAAGTTGAATGAGCCATGCTCTAGTGTGAAGACAG	277	
Qy	242	GATTGGATTATGATATGATTCTGATGACGAAGAAGAGGAGCTTATCAACTCGAAGAG	301	
Dp	278	GATTGGATTATGATATGATTCTGATGACGAAGAAGAGGAGCTTATCAACTCGAAGAG	337	
Qy	302	TGTTCTCATTGAGCCAAAGTGGTAAATCTGGCATTTGATGACCTTCATGTCAGAGAC	361	
Dp	338	TGTTCTCATTGAGCCAAAGTGGTAAATCTGGCATTTGATGACCTTCATGTCAGAGAC	397	
Qy	362	CAGGGATTACAGCTCAATATCAATGCGCTGCTCATGATCTTGAGAGAAATCAT	421	
Dp	398	CAGGGATTACAGCTCAATATCAATGCGCTGCTCATGATCTTGAGAGAAATCAT	457	
Qy	422	TCTATTAGCTTTGAGCTGAGTGTGTTCTGAACATCTGCTTAAGGAATGAAGGTGC	481	
Dp	458	TCTATTAGCTTTGAGCTGAGTGTGTTCTGAACATCTGCTTAAGGAATGAAGGTGC	517	
Qy	482	TGTTCAAGAAGCTGAAGATATTTGGCTTAGAGCCCAATGCTCAATCTCAACATT	541	

Mon May 24 08:18:49 2004

us-09-931-457a-30.rmpb

Page 3

```
Db 518 TGTTCAGAGAGCTGAGAGATATGCTAGAGAGCCCAATGCTACACTTCAAGATT 577
Qy 542 TGAAGAGCCCTGCAATCCCAAGGTTCTATTGAAACCACTGCTCCAGAGATANGAAGG 601
Db 578 TGAAGAGCCCTGCAATCCCAAGGTTCTATTGAAACCACTGCTCCAGAGATANGAAGG 637
Qy 602 CTCGATGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Db 638 CTCGATGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697
Qy 662 TGTCTGAGAAATATCTTAAAGAGCAATCCCAATATTAAGCTGATGATGATGATGATGAT 721
Db 698 TGTCTGAGAAATATCTTAAAGAGCAATCCCAATATTAAGCTGATGATGATGATGATGAT 757
Qy 722 TGAAGAGCTGATGCTCTCAGAGAGAAAGCTGCTGCTCAACAGATTCAGAGATGATGAT 781
Db 758 TGAAGAGCTGATGCTCTCAGAGAGAAAGCTGCTGCTCAACAGATTCAGAGATGATGAT 817
Qy 782 TGTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
Db 818 TGTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877
Qy 842 TGTATGAGCAATGAGAAATGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
Db 878 TGTATGAGCAATGAGAAATGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
Qy 902 ATCTTCGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
Db 938 ATCTTCGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Qy 962 GAAAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
Db 998 GAAAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
Qy 1022 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
Db 1058 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
Qy 1082 CTCTCACTAAGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
Db 1118 CTCTCACTAAGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
Qy 1142 ACTGTTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1201
Db 1178 ACTGTTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1237
Qy 1202 CCCCTGAGAGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
Db 1238 CCCCTGAGAGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
Qy 1262 AACATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1321
Db 1298 AACATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1357
Qy 1322 TGTGATTTCTTGTGTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
Db 1358 TGTGATTTCTTGTGTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1392
```

```
RESULT 3
US-10-424-599-55164
; Sequence 55164, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 265684
; SEQ ID NO 55164
; LENGTH: 3075
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_NRT3847_20824C.1
US-10-424-599-55164

Query Match 48.5%; Score 660.8; DB 13; Length 3075;
Best Local Similarity 77.8%; Pred. No. 6.4e-176;
Matches 797; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

Qy 86 TATATGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 145
Db 308 TATATGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 367
Qy 146 CCATTAATGTAATCTAATTAACCTGCGAGTGTGTGCTCCGAGTTGCTGTAACCT 205
Db 368 TCAATGTAATCTAATTAACCTGCGAGTGTGTGCTCCGAGTTGCTGTAACCT 427
Qy 206 GGAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 265
Db 428 AGAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
Qy 266 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 325
Db 488 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 547
Qy 326 TATATGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 385
Db 548 TATATGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 607
Qy 386 AATGCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 445
Db 608 CATGCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 667
Qy 446 GGTTCGACAGATCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 505
Db 668 AATGCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 727
Qy 506 GGTTCGACAGATCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 565
Db 728 GGTTCGACAGATCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
Qy 566 TCAATTAATGTAATCTAATTAACCTGCGAGTGTGTGCTCCGAGTTGCTGTAACCT 625
Db 788 TCAATTAATGTAATCTAATTAACCTGCGAGTGTGTGCTCCGAGTTGCTGTAACCT 847
Qy 626 TGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
Db 848 TGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907
Qy 686 GAATCCGATTAATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 745
Db 908 GAATCCGATTAATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 967
Qy 746 AAAGCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 805
Db 968 AAAGCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1027
Qy 806 AGTCAATCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 865
Db 1028 AGTCAATCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1087
Qy 866 GCTTCCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 925
Db 1088 ACTTCCTGCTGCTGTAAGAGTCCGAAATGCTCAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1147
Qy 926 TGTCTTGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 985
Db 1148 TGTCTTGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1207
```

QY 986 CAGCTTCGGGAGAGTACTCTCTCCGTCGTAATTGAGTCAAGTCAAGCCGAGCTGA 1045  
DB 1208 AAGCTTGGAGAGGCTTATTCATCTGCTCTTGAATCCGTAAGCAGAGAGAGA 1267  
QY 1046 AAGCATACCTTTAGCCCTGAATTCGCGTTAAGGCTCTCACTAGTAATTTCTGTT 1105  
DB 1268 GAGCATGCTTCGAGCCTTAATTAAGTTAATGAGCATTTGCTGCTTTTCAAGTT 1327  
QY 1106 ACTT 1109  
DB 1328 GCAT 1331

## RESULT 4

US-09-938-842A-893  
Sequence 893, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 893  
LENGTH: 966  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-893

Query Match 43.7%; Score 595.2; DB 9; Length 966;  
Best Local Similarity 76.5%; Pred. No. 1,2e-157;  
Matches 748; Conservative 0; Mismatches 218; Indels 12; Gaps 1;  
QY 90 ATGCTGTTGAAGGTCGGAATGCCAAGATGTTAGCAAGATGTTGTTAAACCCCA 149  
DB 1 ATGCTGTTGAAGGTCGGAATGCCAAGATGTTAGCAAGATGTTGTTAAACCCCA 60  
QY 150 TTGATATATCTAATAAAGCTTGGAGTGTGTTGCTCCGAGTGTGCTGCTAACTGAG 209  
DB 61 TTAGGTATCTAATAAAGCTTGGAGTGTGTTGCTCCGAGTGTGCTGCTAACTGAG 120  
QY 210 TTGATGAGCAATGCTAGTGTGAGAGCAAGATGTTGAGTATGATGATGCTGATGCA 269  
DB 121 ATGATGAGCAATGCTAGTGTGAGAGCAAGATGTTGAGTATGATGATGCTGATGCA 180  
QY 270 GAAAGGAAGGACTTATGACCTGGAAGAGTGTGCTTCACTTGAAGCAAGAGTAT 329  
DB 181 GAAAGGAAGGACTTATGACCTGGAAGAGTGTGCTTCACTTGAAGCAAGAGTAT 240  
QY 330 ACTGCGATTTGATGAGCTTCAAGTCAAGAGGAGTCAAGAGTCAATATTATCAATG 389  
DB 241 ACTGCGATTTGATGAGCTTCAAGTCAAGAGGAGTCAAGAGTCAATATTATCAATG 300  
QY 390 CCTGCTTTATGAGCTTGAAGAGAAATCAATTTAGCTTTTGAAGCTGAGTTGTT 449  
DB 301 CCAAGCTTCAATGAGTATGAGAGAGAAATCAATTTAGCTTTTGAAGCTGAGTTGTT 360  
QY 450 CTGACAGATCTGCTAGAGAGATGAAAGTGTGTTGAGAGGCTGAGAGATTTGAGCT 509  
DB 361 CTGACAGATCTGCTAGAGAGATGAAAGTGTGTTGAGAGGCTGAGAGATTTGAGCT 420  
QY 510 AAGAGCCCAATGCTTACATCTTCAACAAATTGAAACCTGCGCAATCCCAAGGTTAT 569

DB 421 AAAACACCTAATGTTATATGCTTCAACAGTTTGAAATCTGCTAATCCAAAGATCCAT 480  
QY 570 TATGAACACCTGATGCTCAAGATATGAAAGGCTCGATGGAATATGATGATTTGTT 629  
DB 481 TATGAGATCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 630 TCTGGATGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 689  
DB 541 TCTGGATGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
QY 690 CCAATATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 749  
DB 589 ACAACATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 648  
QY 750 CCGTGTCCACAAAGATTCAGAGATTTGCTGCTGTTTATCCCTGATGCTTGAAGTC 809  
DB 649 CCAAGTCCCAAAATTCAGAGATTTGCTGCTGTTTATCCCTGATGCTTGAAGTC 708  
QY 810 AATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 869  
DB 709 GATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768  
QY 870 CTGCGCTTAAAGAGGCTTATTTGAGATATCTTCCGAGCTGAGCTGCTGCT 929  
DB 769 CTGCTAGGAGAGAGGCTTATTTGAGATATCTTCCGAGCTGAGCTGCTGCT 828  
QY 930 TTTCAGATTCGAAAG 989  
DB 829 ATCAACATTCGAAAG 888  
QY 990 TTGCGGAG 1049  
DB 889 TTGCGGAG 948  
QY 1050 ATGACTTTGAGAGCTGCA 1067  
DB 949 ATGACTTTGAGAGCTGCA 966

## RESULT 5

US-09-938-842A-893  
Sequence 893, Application US/09938842A  
Patent No. US20040009476A9  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 893  
LENGTH: 966  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-893

Query Match 43.7%; Score 595.2; DB 11; Length 966;  
Best Local Similarity 76.5%; Pred. No. 1,2e-157;  
Matches 748; Conservative 0; Mismatches 218; Indels 12; Gaps 1;  
QY 90 ATGCTGTTGAAGGTCGGAATGCCAAGATGTTAGCAAGATGTTGTTAAACCCCA 149

Db 1 AAGGCTTCAGTGGCTCCCTAAATGCTGTAAGATGTAACGTAATAGGAACTCA 60  
 QY 150 TATGATATCTAAATTAACCTGCGAGTGGTGTGTCGCCGGTGTCTCTAACTGAG 209  
 Db 61 TATGATATCTAAATTAACCTGCGAGTGGTGTGTCGCCGGTGTCTCTAACTGAG 120  
 QY 210 TATGATATCTAAATTAACCTGCGAGTGGTGTGTCGCCGGTGTCTCTAACTGAG 269  
 Db 121 AAGTATGAAACCAATGCTTACGCTGCAAGACAGATGTTATGATGATGCTATGCA 180  
 QY 270 GAAGAAAGGAGCTTATCACTGCAAGTGAAGAGTGTCTCTATGAGGCACTAAGTGTAT 329  
 Db 181 GAAGCTAAAGAGCTTATCACTGCAAGTGAAGAGTGTGTGATGATGAGCACTAAGTGTAT 240  
 QY 330 ACTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389  
 Db 241 ACTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 390 CCGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449  
 Db 301 CCGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 QY 450 CTGACGATCTGCTGTAAGGAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 509  
 Db 361 CTGACGATCTGCTGTAAGGAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
 QY 510 AAGAGGCGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 569  
 Db 421 AAGAGGCGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 480  
 QY 570 TATGAAACCACTGCTGCAAGATGATGAAAGGCTCCGATGAGAAATGATGATGATGAT 629  
 Db 481 TATGAAACCACTGCTGCAAGATGATGAAAGGCTCCGATGAGAAATGATGATGATGAT 540  
 QY 630 TCTGGATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689  
 Db 541 TCTGGATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588  
 QY 690 CCGAATATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749  
 Db 589 CCGAATATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648  
 QY 750 CCGTGTCCACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809  
 Db 649 CCGTGTCCACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
 QY 810 AATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869  
 Db 709 AATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768  
 QY 870 CTTGCGCTTAAAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929  
 Db 769 CTTGCGCTTAAAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828  
 QY 930 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989  
 Db 829 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
 QY 990 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049  
 Db 889 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
 QY 1050 ATGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1067  
 Db 949 ATGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966

RESULT 6  
 US-10-424-599-55160  
 ; Sequence 55160, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; NUMBER OF SEQ ID NOS: 2003-04-28  
 ; SEQ ID NO: 55160  
 ; LENGTH: 1404  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_20820C.1  
 US-10-424-599-55160

Query Match 42.5%; Score 578.6; DB 13; Length 1404;  
 Best Local Similarity 74.0%; Pred. No. 7.6e-153;  
 Matches 759; Conservative 0; Mismatches 264; Indels 2; Gaps 2;

QY 86 TATATGCTGCTGTAAGAGTCCGAGATGCTGCAAGATGTTACGATGATGATGATGATGATGAT 145  
 Db 147 TATATGCTGCTGTAAGAGTCCGAGATGCTGCAAGATGTTACGATGATGATGATGATGATGAT 206  
 QY 146 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204  
 Db 207 ACTGACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 266  
 QY 205 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 264  
 Db 267 TAGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 326  
 QY 265 ATGCGAAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 324  
 Db 327 ATGCGAAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386  
 QY 325 GTAATATGCTGCTGTAAGAGTCCGAGATGCTGCAAGATGTTACGATGATGATGATGATGATGAT 384  
 Db 387 GAAACAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446  
 QY 385 CAAATGCTGCTGTAAGAGTCCGAGATGCTGCAAGATGTTACGATGATGATGATGATGATGATGAT 444  
 Db 447 CAAATGCTGCTGTAAGAGTCCGAGATGCTGCAAGATGTTACGATGATGATGATGATGATGATGAT 506  
 QY 445 TGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504  
 Db 507 TGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566  
 QY 505 TGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564  
 Db 567 TGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626  
 QY 565 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 624  
 Db 627 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 685  
 QY 625 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 684  
 Db 686 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 745  
 QY 685 AAGATTCGATATTAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744  
 Db 746 AAGATTCGATATTAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 805  
 QY 745 GAAAGCTGCTGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 804  
 Db 806 GAAAGCTGCTGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865  
 QY 805 AAGTCAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 864  
 Db 866 AAGTCAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925

QY 865 AGCTTCTGCGCTTAAAGAGGCTATTGTTGGAAATATCTCCGAGCTCAAGCTGCTG 924  
DB 926 AGCTTCTGCAATGAAAGAGGCTTACTGTTGGAAATATCTGCTGCTGCTGAGCTG 985  
QY 925 CTGCTTTGAGATTGCAAAAAGACCAAAAATGCCGAGAGCTTATTGTTGCCCTTTTC 984  
DB 986 CTGCAATTGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1045  
QY 985 CCAAGCTTGGGAGAGGAGTACCTGCTCCGCTGCTTATTGAGTCAAGTGAAGAGAGCTG 1044  
DB 1046 CAAAGTTTGGAGAGGCTTATCTATGCTGCTGCTTGAATCCCTGAAGAGAGAGAG 1105  
QY 1045 AAAGCATGACTTTGAGCCCTGGAATTCCTGTTAAAGCTCTCACTAATTTTCTTGT 1104  
DB 1106 AGAGCTTGGCTTTTGAAGCTTGAATGATGATGATGATGATGATGATGATGATGAT 1165  
QY 1105 TACTT 1109  
DB 1166 TGCAT 1170

## RESULT 7

US-10-425-114-25435  
Sequence 25435, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 25435  
LENGTH: 1319  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3900-005-C7\_FLI  
US-10-425-114-25435

Query Match 40.9%; Score 557.4; DB 13; Length 1319;  
Best Local Similarity 72.7%; Pred. No. 7.5e-147;  
Matches 720; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 75 TAGTGAAGAGATATATATGCTGTTGAAGTCCGGAATGCAAGATGTTACGGAATG 134  
DB 76 TCGTGAAGAGATATATATGCTGTTGAAGTCCGGAATGCAAGATGTTACGGAATG 135  
QY 135 ATGCTAAACCCCAATGAT 194  
DB 136 ATGCTAAACCCCAATGAT 195  
QY 195 GCTGTAACCTGAGTGTATGAGAGCACTCTATGTTGAAGAGAGAGAGAGAGAGAGAGAG 254  
DB 196 GCTGTAACCTGAGTGTATGAGAGCACTCTATGTTGAAGAGAGAGAGAGAGAGAGAGAG 255  
QY 255 ATGATGCTGATGAG 314  
DB 256 ATGATGCTGATGAG 315  
QY 315 CCAAGAGAGATATATATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 374  
DB 316 CCAAGAGAGATATATATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 375  
QY 375 CTGATATATATATATATATGCTGTTGATGAGATGAGATGAGATGAGATGAGATGAGAT 434  
DB 376 CTGATATATATATATATGCTGTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 435

QY 435 GAGCTGAGTGGTCTTGAAGATGCTGCTTAAAGGAAATGAAGAGTGTCTTCAAGAGCT 494  
DB 436 GGTGCTGATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 495  
QY 495 GAAAGATATATGCTTAAAG 554  
DB 496 GAAAGATATATGCTTAAAG 555  
QY 555 AATCCAGAGTATATATGAAACCACTGCTGCAAGATATGAGAGAGAGAGAGAGAGAGAG 614  
DB 556 AATCCAGAGTATATGAAACCACTGCTGCAAGATATGAGAGAGAGAGAGAGAGAGAGAG 615  
QY 615 ATGATGATATGTTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674  
DB 616 ATGATGATATGTTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 675  
QY 675 CTTAAAG 734  
DB 676 CTTAAAG 735  
QY 735 CTCTCAAG 794  
DB 736 TTGAATGATGAG 795  
QY 795 GGTGCTTGAAGATCAATCTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854  
DB 796 GGTGCTTGAAGATCAATCTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855  
QY 855 GAACTGCAAG 914  
DB 856 GAACTGCAAG 915  
QY 915 GCACTGCTGCTGCTTCTTCAAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 974  
DB 916 GCTGCACTGCTGCTGCTTCTTCAAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 975  
QY 975 GCGTTTTTCCAGCTTCCGAG 1034  
DB 976 GCGTTTTTCCAGCTTCCGAG 1035  
QY 1035 CCGAG 1065  
DB 1036 AAG 1066

## RESULT 8

US-10-425-114-19467  
Sequence 19467, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 19467  
LENGTH: 1308  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3150-027-D8\_FLI  
US-10-425-114-19467

Query Match 40.8%; Score 555.8; DB 13; Length 1308;  
Best Local Similarity 72.6%; Pred. No. 2.1e-146;

Matches 719; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

QY 75 TAGTGGAGCATATATATGCTGTGTAAGAGTCCGGAATTCGAAAGATGTAACGAAATG 134
DB 65 TGTGACGTGACCAATGAGAGAGGCTCGCTGATGACCAAGACCTCAACCAATG 124
QY 135 ATTGTGTAACCCCATTAATATCTAAATTAACCTGCGAGATGTTGTGTCGCGGCT 194
DB 125 ATGTGGAACACGCGCTGTGTATCTCAACAAAGTACGATGGTGTGCTGCGCGCTC 184
QY 195 GCTGTAAACTGAGTGTGATGAGACCATCTCTAGTGTAGAGACAGATTTGGTATAGT 254
DB 185 GCGGCAACCTGAGTGTGATGAGAGGCTCTCTCAAGGTCAAGGATGATGGCTACAGC 244
QY 255 ATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 314
DB 245 ATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY 315 CCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 374
DB 305 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 364
QY 375 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 434
DB 365 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 424
QY 435 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 494
DB 425 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
QY 495 GAAGATATATGCTTAAGACGCGCAATGCTTACATCTTCAACATTTGAAACCTGTC 554
DB 485 GAAGATATATGCTTAAGACGCGCAATGCTTACATCTTCAACATTTGAAACCTGTC 544
QY 555 AATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 614
DB 545 AATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604
QY 615 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 674
DB 605 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 664
QY 675 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 734
DB 665 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 724
QY 735 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 794
DB 725 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 784
QY 795 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 854
DB 785 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
QY 855 GAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 914
DB 845 GAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
QY 915 GCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 974
DB 905 GCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 964
QY 975 GCGGTTTTCGAGCTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATG 1034
DB 965 GCGGTTTTCGAGCTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATG 1024
QY 1035 CGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065
DB 1025 AAGGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1055

```

RESULT 9  
US-10-425-114-23298

Sequence 23298, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalev, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 23298  
LENGTH: 1312  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB597-017-F4\_FLI  
US-10-425-114-23298

Query Match 40.8%; Score 555.8; DB 13; Length 1312;  
Best Local Similarity 72.6%; Pred. No. 2,1e-146;  
Matches 719; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

QY 75 TAGTGGAGCATATATATGCTGTGTAAGAGTCCGGAATTCGAAAGATGTAACGAAATG 134
DB 69 TGTGACGTGACCAATGAGAGAGGCTCGCTGATGACCAAGACCTCAACCAATG 128
QY 135 ATTGTGTAACCCCATTAATATCTAAATTAACCTGCGAGATGTTGTGTCGCGGCT 194
DB 129 ATGTGGAACACGCGCTGTGTATCTCAACAAAGTACGATGGTGTGCTGCGCGCTC 188
QY 195 GCTGTAAACTGAGTGTGATGAGACCATCTCTAGTGTAGAGACAGATTTGGTATAGT 254
DB 189 GCGGCAACCTGAGTGTGATGAGAGGCTCTCTCAAGGTCAAGGATGATGGCTACAGC 248
QY 255 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 314
DB 249 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
QY 315 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 374
DB 309 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 368
QY 375 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 434
DB 369 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 428
QY 435 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 494
DB 429 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 488
QY 495 GAAGATATATGCTTAAGACGCGCAATGCTTACATCTTCAACATTTGAAACCTGTC 554
DB 489 GAAGATATATGCTTAAGACGCGCAATGCTTACATCTTCAACATTTGAAACCTGTC 548
QY 555 AATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 614
DB 549 AATCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 608
QY 615 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 674
DB 609 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
QY 675 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 734
DB 669 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 728
QY 735 CTCTAATTAATCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 794

```



Db	729	TTGATGATGAGAAACCTGGACCAACAAGATTCAAGAAATGGAGCTGTTTATCCCT	788
Qy	795	GGTCTCTTGGAAAGTCATCTTCTTGAATGATGTCAAATATCAAGTGAAGAAACATA	854
Db	789	GGATCTTTGAGATGTGATCTCTCTGATATAAATCTACAGGTTTCAAGTGAAGAAAGTATTC	848
Qy	855	GAACCTGCAAGCTCTTCGCGCTTAAAGAGGCCATTTGTGGGAAATCTTCCGAGCT	914
Db	849	GAGACTGGCAAGGCTCTTGCTCTGTAAGAAAGGGTTGTGTGGAATCTCTTGATGCA	908
Qy	915	GCACCTGCTGCTCTTTCAGATTTGGCAAAAAGACAGAAATATCCGGGAACCTATATGTT	974
Db	909	GCTCAGCTGCGCGCATGATGCTGTCTATAGAGGCCAGAAACGCCGGAAAGCTATTTGTT	968
Qy	975	GCCGTTTTTCCACAGCTTCGGGGAGAGGTAACCTGCTCCCTCGTGTCTATTTGAGTCAGTGA	1034
Db	969	GTCGTGTTCCGAGAGCTTCGGCGAAGGCTACCTCTATCGGTGCTGTTCATTCATCAAG	1028
Qy	1035	CGGAGAGCTAAAGCATATCTTTGAGGCTT	1065
Db	1029	AAGAAAGCGAGAGCATATGTGTCAAGCCCT	1059

RESULT 10  
US-10-425-114-1445  
; Sequence 1445, Application US/10425114

```

1 Publication No.: US20040034888A1
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Liu, Jindong
6
7 APPLICANT: Zhou, Yihua
8
9 APPLICANT: Kovalic, David K.
10
11 APPLICANT: Screen, Steven B.
12
13 APPLICANT: Tabaska, Jack E
14
15 APPLICANT: Cao, Yongwei
16
17 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
18
19 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
20
21 FILE REFERENCE: 38-21(5313)B
22
23 CURRENT APPLICATION NUMBER: US/10/425,114
24
25 NUMBER OF SEQ ID NOS: 73128
26
27 SEQ ID NO 145
28
29 LENGTH: 1403
30
31 TYPE: DNA
32
33 ORGANISM: Zea mays
34
35 FEATURE:
36
37 OTHER INFORMATION: Clone ID: 700153393_FLI
38
39 US-10-425-114-1445

```

Query Match	40.8%	Score	555.8	DB 13	Length	1403			
Best Local Similarity	72.6%	Pred. No.	2.2e-146						
Matches	719	Conservative	0	Mismatches	272	Indels	0	Gaps	0

Qy	75	TAGGAGCAGATATATATGGCTGTGAAAGAGTCGGATATGCAAGAAGTTACGGAATG	134
Db	160	TGTGACGTGACCCCATGGAGAGGACCTGCGGTGATCGCAAGAGCGTCAACGAGTTG	219
Qy	135	ATTGTAAACCCCATTTAGTATATCTAATAATCTGGGATGTGTGTGTCGCCGGTT	194
Db	220	ATCGGAAACAGCGCGCTGTGTACTCTCAACAGGTGACCGATGGGTGCGTGGCGGCTTC	279
Qy	195	GCTCTAAACTGAGATTGATGGAGCCATGCTCTAGTGAAGACACGAGATTGGGTATAGT	254
Db	280	GCCGCCAAGCTCGAGTCCATGGAGCCGTGCTCCAGCGTCAAGGATAGGATTGGCTACAGC	339
Qy	255	ATGATTGTCTATGCAAGAGAGAGGGAATTATCACACTGGAAGAGTCTCTCATTAG	314
Db	340	ATGATCAACGACGACAGAGAGAAAGGCGTGATTACTCCAGCGTGAAGTGTCTGATTTAA	399
Qy	315	CCAACAAAGTGTAACTACTGGAATTGAGATTAGCTTCATGGACGACGCCAGGGTTACAAG	374
Db	400	CCAATACGGCAACACAGAGCATTTGACTTGCGCTTTATGGCTGCTGCAAGGGCTACAAA	459
Qy	375	CTCATTAATACAATGCGCTGCTCTATAGAGTCTTGAGAAAGATATTCTATTAGCTTT	434

Db	460	CTTACATCTCAAAATGCCCTCCCTCCATGAGCATGAGAGAGGATCATATATGAAGCTTTT	519
Qy	435	GGAGCTGAGTTGGTCTTGACAGATCTCTGTAAGGAAATGAAAGTGCTTTCAAGAGCT	494
Db	520	GGTCTGAACTTGTCTTTACTGACCCACTCCCTGGGAAATGAAAGAGCTGTCAAGAAAGG	579
Qy	495	GAAGAGATATGGCTTAAGACGCCCAATGCTTACACTTCAACAAATTTGAAAAACCTTGC	554
Db	580	GAAAGATACAAAGCAAAAGACACCCAACTGTATACATCTTTCAACAAATTTGAAATCCAGCT	639
Qy	555	AATCCCAAGTTCAATTATGAAACCACTGCTCAGAGATATGAAAGGCTCCGATGGAAA	614
Db	640	AACCCAAGATTCATCTATGAGACTATCTGGCTGTAAATCTGAAAGACTACAGAGGAAA	699
Qy	615	ATTGATGCATTTGTTTCTGGATAGGCACTGGTGTACATPAAAGTCTCGAAAAAT	674
Db	700	ATTCTGSCCTTGTATCTGTATCGGAGAGAGGTACATCACCGTACTGGGCGATAC	759
Qy	675	CTTAAAGGCGAATCCGAAATATAAAGCTGATGCTGTGGAACCAAGTGAAGTCCAGTG	734
Db	760	CTCAGAGGCAAAATCTTAATGTCAAGTCTATGTGTGAGACCAGTGGAGATGCTGTT	819
Qy	735	CTCTCAGAGGAAAGCCTGTCTCAACAAGATTCAAGGATTGCTGCTGTTTATCCCT	794
Db	820	TTGATGATGTGAAAACTGGACACACAAGAATTCAAGGAATGGAGAGCTGTATTACCT	879
Qy	795	GGTGTCTTGAAGTCAATCTTCTTGAATGAAGTGTCAAAATCAAGGATGAGCAATA	854
Db	880	GGAAGCTTGAAGTGTATCTCTCTGATAGACTCTACAGGTTTCAAGTATGAAGTATT	939
Qy	855	GAACCTGCAAACTTCTTGCGCTTAAAGAGGCTATTGTGGAATACTTCCGAGCT	914
Db	940	GAGACTGTCAGAGCTCTTGCTCTGAAATAAGGATGTGTGTAATCTTCTGTGTGA	999
Qy	915	GCACCTGTGTGCTTTTCAGATGCAAAAAAGACCAAGAAATCCGGGGAAGCTATGTT	974
Db	1000	GCTGAGTGTGGGCAATTAGGCTGTCTTAAAGGCCAAGAAAACGCCGGAAGCTATTTGTT	1059
Qy	975	GCCGTTTTTCCCAAGTTGCGGGAAGAGTACCTGTCTCCGTGCTATTTGAGTGCAGTGA	1034
Db	1060	GTCGTGTTCCCGAAGTTGCGGAGAGGCTACACTCTCATCGTGTCTTCCAGTCCATCAG	1119
Qy	1035	CGCGAAGCTAAAGCATGACTTTTGAGTCT	1065
Db	1120	AAAGAAAGGAAAGCATGTGTGAGGCTCT	1150

## RESULT 11

US-10-425-114-23608  
; Sequence 23608, Application US/10425114  
; Publication No. US20040034888A1

```

1  GENERAL INFORMATION:
2  APPLICANT: Liu, Jindong
3  APPLICANT: Zhou, Yihua
4  APPLICANT: Kovalic, David R.
5  APPLICANT: Screen, Steven E
6  APPLICANT: Tabaska, Jack E
7  APPLICANT: Cao, Yongwei
8  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
9  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
10 FILE REFERENCE: 38-21(5313)B
11 CURRENT APPLICATION NUMBER: US/10/425.114
12 CURRENT FILING DATE: 2003-04-28
13 NUMBER OF SEQ ID NOS: 73128
14 SEQ ID NO 23608
15 LENGTH: 1301
16 TYPE: DNA
17 ORGANISM: Zea mays
18 FEATURE:
19 OTHER INFORMATION: Clone ID: LIB3597-067-G10_F11
20 US-10-425-114-23608

```



Query Match 40.7%; Score 554.2; DB 13; Length 1301;  
Best Local Similarity 72.5%; Pred. No. 6e-146;  
Matches 718; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

75 TATGAGCAGATTAATATGCTGTTAAAGGTCGGAATGCGCAAGATGTAAGATG 134  
58 TGTGACGTGACCCATGGAAGAGGCTGCTCCATGATCCAGAGACCTGACAGATG 117  
135 ATTGTAAGAACCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 194  
118 ATCGGGAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177  
195 GCTGCTAACTGAGTGAATGAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 254  
178 GCGCGAGCTGAGTGAATGAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 237  
255 ATGATGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 314  
238 ATGATGACGAG 297  
315 CCAACAGAGTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374  
298 CCAACTAG 357  
375 CTGATTAATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434  
358 CTGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417  
435 GAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494  
418 GGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477  
495 GAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554  
478 GAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537  
555 AATCCAGAGTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614  
538 AATCCAGAGTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597  
615 ATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 674  
598 ATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657  
675 CTTAAGAGCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734  
658 CTTAAGAGCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
735 CTTAAGAGCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794  
718 TTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777  
795 GGTGCTGAGTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854  
778 GGTGCTGAGTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837  
855 GAACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914  
838 GAACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897  
915 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974  
898 GCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957  
975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034  
958 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017  
1035 GCGGAGCTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065  
1018 AAGGAGCTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048

RESULT 12  
US-10-424-599-61234  
Sequence 61234, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongsui  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 61234  
LENGTH: 1603  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_26307C.1  
US-10-424-599-61234

Query Match 39.9%; Score 543.4; DB 13; Length 1603;  
Best Local Similarity 73.1%; Pred. No. 7.9e-143;  
Matches 697; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

108 GGAATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167  
233 GGAATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292  
168 CTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227  
293 ATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 352  
228 ACTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287  
353 ACTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412  
288 ACACCTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347  
413 ACTGGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472  
348 TTGATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407  
473 TTGATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532  
408 GAGAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467  
533 GAGAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592  
468 GGAATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527  
533 GGAATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652  
528 AATCTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587  
653 AATCTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712  
588 GAGATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 647  
713 GAGATGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772  
648 GGTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707  
773 GGTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832  
708 GGTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767  
833 GGTGCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892  
768 CAAGGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827

Mon May 24 08:18:49 2004

us-09-931-457a-30.rnpb

Page 10

Db	CAAGGAATTGGTGGTATCATCTCCGAAAGTTGGATGTAAATCTCTGATGAAGTT	952
Qy	GTTCAAAATATCAAGTATGAAGCAATACAAACTGAAGCTTTGGCTTAAGAAAGG	887
Db	ATTACAGTTTAAAGTAAGAGGCTATAGAAATGCTAAACGTGGCTTGAAAGAGT	1012
Qy	CTATTTCGGGAATATCTTCGGAAGCTCAGCTGCTGCTTTTCATTTGCAAAAGA	947
Db	TTCGTGATGGGAATTTTCATCAGGAGCTGTGACGACGACAAATAAAGTGGGAAGGA	1072
Qy	CCAGAAATGCCGGGAAGCTTATTTGTTCGGTTTTCCCAAGTTCCGGGAAGGTAAGTCTG	1007
Db	CCGAGAAATGACGAAAGCTCATTTGTGATTTTCCCAAGTTTGGAAAGCTTACTTA	1132
Qy	TCTTCCTGCTATTTGAGTCAGTGAAGCGAGAGCTGAAGCATATCTTTTGA	1060
Db	TCTTCACCACTGTTTGATTCATTAAGAAAGAGCTGAACAAATGACATTTGA	1185

```

1      RESULT 13
2      US-10-425-114-22525
3      Sequence 22525, Application US/10425114
4      Publication No. US2004003488A1
5      GENERAL INFORMATION:
6      APPLICANT: Liu, Jingsong
7      APPLICANT: Zhou, Yihua
8      APPLICANT: Kovalic, David K.
9      APPLICANT: Screen, Steven E
10     APPLICANT: Tabaska, Jack E
11     APPLICANT: Cao, Yongwei
12     TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
13     TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
14     FILE REFERENCE: 38-21(53313)B
15     CURRENT APPLICATION NUMBER: US/10/425,114
16     CURRENT FILING DATE: 2003-04-28
17     NUMBER OF SEQ ID NOS: 73128
18     SEQ ID NO 22525
19     LENGTH: 1279
20     TYPE: DNA
21     ORGANISM: Zea mays
22     FEATURE:
23     OTHER INFORMATION: Clone ID: LIB587-229-H7_F1
24     US-10-425-114-22525

```

Query Match	Similarity	38.5%	Score 524.4	DB 13	Length 1279
Best Local	Similarity	74.2%	Pred. No. 1.6e-137		
Matches	Conservative	0	Mismatches 231	Indels 0	Gaps 0
QY	172	CGAGTGGTGTGTGTCCCGGGTGTGCTGTAACCTGAAGTTGATGAGCCATGCTCTAGTG	231		
Dh	133	CGAGTGGTGTGTGTCCCGGGTGTGCTGTAACCTGAAGTTGATGAGCCATGCTCTAGTG	192		
QY	232	TGAAGGACAGAGATTGGGGTATAGTATGATTTGCTGATTCGACGAAAGAGAGACTTATCAAG	291		
Dh	193	TGAAGGATGAGATTGGCTTACAGCAAGTATCAAGAGCGAGAGAGAGAAAGGCTGATTACTC	252		
QY	292	CTGGAAAAGTGTGCTCATTTGAGCCAAACAAGTGTAAATCTGGCATTTGATTAAGCTTCA	351		
Dh	253	CAGGGTATGATGTCTTGATTTGAACCAACTAGGGGACACAGGATTTGACCTGGGCTTTA	312		
QY	352	TGGCAGCAGCCAGGGGTTACAAAGCTCATATTAATCAATGCTGCTTCTATAGTCTTTGAGA	411		
Dh	313	TGGCTGCTGCCAAGGGGCTACAAACTTACACTCAATGCTGCTGCTCCCTCCACATGACATGGAGA	372		
QY	412	GAGGATATTTCTATTAGCTTTTGAAGCTGATTTGGTCTGACAGATCCTGTAAAGGAA	471		
Dh	373	GAGGATATATTTGAAGGCTTTTGGTGTGTAACCTTCTTACTGACCACTTCTGGGAA	432		
QY	472	TGAAGTGTCTGTTCAGAGGCTGAGAAGATTTGGCTAAGACGCCCAATGCTTACATAC	531		
Dh	433	TGAAGGAGCTGTCAAGAAAGCGAAGAGATTCAGAGCAAAAGACACCCAACTGTACATCC	492		
QY	532	TTCAACAATTTGAAAAACCTCGCAATCCCAAGTTCATTTATGAACCACTGCTCCAGAGA	591		

Db	493	TTCAACAAATTGAAATCCAGCTAACCCAAAGATTCACATATGAGACATCTACGCGCTGAAA	552
Qy	592	TATGAAAAGCGTCCCATGGGAAAATGATGCAATTGTTCTTCGGATPAGGCATCGTGTGTA	651
Db	553	TCGGAAGGCTACAGAGGAAAATGCTGCGCTTGATATCTGTATGCGTGAACAGAGGTA	612
Qy	652	CAATAAACAGTGTGTGAAAATATCTTAAAGGCGAATCCGATATATAAGCTGATGTTGTG	711
Db	613	CCATTCACCGGTACTGGGCGATACCTCAGAGGCAAAAATCCTAAATGTCMAACTATATGTTG	672
Qy	712	TGAAACAGTGAAGTCCAGTGTCTCAGAGGAAAAGCTGTGTCCACAGAATTCAG	771
Db	673	TGAGCGCATGTAGAGATGTCTTTTGAATGTGTGAAAACCTGACACACACAAAGTTCAAG	732
Qy	772	GAATTGTGTGTGTTTATCCCTGTGTCTTGGAAATCAATCTTCTTGAATGAAGTTCTC	831
Db	733	GAATTGAACTGTGTTTATCCCTGAGGCTTGGATGTGTGATCTCTTGAATGAACCTTAC	792
Qy	832	AAATATCAAGTGTGACGATPAGAAATCTGCAAAAGCTCTTCGCGCTTAAAGAAAGGCTAT	891
Db	793	AGGTTTCAAGTATATGAGCTATTTGAGCATGTGCAGAGGCTCTTGCTGTAAGAAAGGCTGT	852
Qy	892	TTGTGGAAATATCTTCGGAAGCTGCAGCTGTCTGTCTTTCAATTTGCAAAAAGACAG	951
Db	853	TGATTTGAAATCTCTTCTGTGTGACGTGCAGCTGCGGACGTGAGGCTTGTCAAGAGGCGAG	912
Qy	952	AAATATCCGAGAAAGCTTATTTGTTCGCTTTTCCAGCTTCGGGAGAGGATCCTGTCTT	1011
Db	913	AAAACCCGAGAAAGCTATTTGTGTGCGTTCGCCAGCTTCGGGAGACGCTTACCTCTCAT	972
Qy	1012	CCGTGTATTTAGTCAAGTGAACGACGCAAGGCTGAAGAAGTACCTTTGAGCGCT	1065
Db	973	CGGTGTCTTCCAGTTCATCAGAAAGAGGCGGAAAGCATGTGTGTGAGGCTT	1026

```

RESULT 14
US-10-425-114-34372
; Sequence 34372, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34372
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17258D02_FU1
US-10-425-114-34372

```

	Query Match	38.3%	Score 521.8	DB 13	Length 1490
	Best Local Similarity	71.6%	Pred. No. 9.9e-137		
	Matches	685	Conservative	0	Mismatches 272; Indels 0; Gaps 0;
Oy	111 ATGCCAAGATGTTACGAATTGATGGTAAACCCTATGATTAATAACT	170			
Db	240 ATGCCAACGACGTACCCCACTCATCGGCACACACCAATGGTGATCTCAACAGTC	299			
Oy	171 GGCGATGGTGTGTCGCCGGGTGCTCTAACTGAGTAGATGAGCACAAGCTCTAAGT	230			
Db	300 GTCAAGGATCTGTGGCAATGTCGTCTAAGTCGAGATTATGAGCCCCGTGTAGC	359			

Mon May 24 08:18:49 2004

us-09-931-457a-30.rnpb

Page 11

QY	231	GTGAAGACAGAGATGGGGTATAGTATGATGCTGAGCAGAAGAGAGAAGGACCTTATACA	299
Db	360	GTGAAGACAGAGATGGGGTACAGTATGATATGATATGCTGAGACAGAGGGCTTGATTA	419
QY	291	CTGGAAAGAGTGTCTCATTTAGACCAACAGTGTATATCTGSCATTGATTAAGCTTC	350
Db	420	CTTGAAGAAGGTGTTTGGTGGAAAGCAACAAGTGGAAACAAGGATTTGGTCTTGCTTC	479
QY	351	ATGGACACAGACAGGGGTTTCAAGCTCCATATTAACAATGCCGCTTCTATAGTCTTGAG	410
Db	480	ATTCTCTCTTTAAAGGATATTAAGCTGATCTPAACATGCTCTTCAATATAGATGAGT	539
QY	411	AGAAGATCATTTCTATTAAGCTTTTGGAGCTGAGTGTGTTCTGACAGATCCGTCTAAGGGA	470
Db	540	AGAAGATCTCTCCACAGCTTTTGGTGTGAACTGTTCCTTACAGATGCTGCAAAAGGG	599
QY	471	ATGAAGGTGCTGTTCCAGAGGCTGAAGAGATTTGGCTPAAGACGCCCATATGCTCAATA	530
Db	600	ATGAAGGGGCTTAAAGGTAAAGGCTPAAGAGATTTTAAACAAGACCCAAATTTCTTAATG	659
QY	531	CTTAAACAATTTGAAAAACCTGCAATCCCAAGTTCATTTATGAACAACATGCTGCCAG	590
Db	660	CTTPAACAGTTTCATTAACCTTGCCAACTTPAAGTTCATTTAGAGATCTACTGCTCCAG	719
QY	591	ATATGAAGAAGCTCCAGTGGGAAATTTGATGATTTGTTTCGAGTATGACACTGTGAGT	650
Db	720	ATCTGGAGAGATTCAAAGGGAAAGTGAATATTTCAATGGTGGAAATGGAAACAGGGGGG	779
QY	651	ACATATACAGTGTCTGGAATAATCTTAAAGACAGAAATCCGAAATTTAAAGCTATTTGT	710
Db	780	ACATATCTGTGTCCCGCCGTTTTCACAGAGAAAAATCCGGAATTTAAGGTATTTGT	839
QY	711	GTGAACCAAGTGTAAAGTCCAGTGTCTCAGAGAGAAAGCTGTGTCACAACAATTCMA	770
Db	840	ATTGAGCTTCTGAAGTAAACATCTCTCCGTGGGAAACCTGTGTCACAATAATCCAG	899
QY	771	GGGATGGTGTCTGTATTAATCCCTGGTGTCTTGAAGTCATCTTCTGATGAAGTTGT	830
Db	900	GGATGGCCAGAGATTTGTTCCAGAGAACTTGGATACGATATTTCTTGAAGATTAAT	959
QY	831	CAATATACAGTATGAAGCAATAGAAATCGAAAGCTTCTTGCGCTTAAGAAGGCTA	890
Db	960	GAGATACCAAGTATGAAGCTGTGGAACAAGAAACAGTGGCTGTCCAGAAAGATTA	1010
QY	891	TTTGTGGGAATATCTTCCGAGCTGACGTGCTGTGCTTTTCAATGTCAAAAAGACA	950
Db	1020	CTGTGTGGAAATCTCTCTGTGGGGCAGAGCCGCCGCCCATTAAGTTGGCCAAAGACA	1070
QY	951	GAAAATGCCGGGAAGCTTATGTTGCCGTTTTCAGCTTTCGGGAGAGGTACTGTCC	1010
Db	1080	GAGAAATGCTGGAAGCTGATAGTGTCCGTGTTCCAGCTTTCGGGAGAGGTACTTTCA	1130
QY	1011	TTCGTGCTATTTAGTCAAGTGAAGCCGCAAGCTGAAGAAGATGACCTTTTGACCTGTA	1067
Db	1140	TCTGTCTCTTATAGTCAATAGAGAGAGATGGGAACATGCACTGTAGCTATGA	1196

RESULT 15  
 US-10-425-114-34527  
 : Sequence 34527, Application US/10425114  
 : Publication No. US20040034888A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Liu, Jingdong  
 : APPLICANT: Zhou, Yihua  
 : APPLICANT: Kovalic, David K.  
 : APPLICANT: Screen, Steven E.  
 : APPLICANT: Tabaska, Jack E  
 : APPLICANT: Cao, Yongwei  
 :  
 : TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 : TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 : FILE REFERENCE: 38-21(53313)B  
 : CURRENT APPLICATION NUMBER: US/10/425,114  
 : CURRENT FILING DATE: 2003-04-28  
 :

```

: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 34527
: LENGTH: 1491
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURES:
: OTHER INFORMATION: clone ID: UC-ZMFLMO17281G01_FLI
US-10-425-114-34527

```

	Query Match	38.3%;	Score 521.8;	DB 13;	Length 1491;
	Best Local Similarity	71.6%;	Pred. No. 9;	96-137;	
	Matches 685;	Conservative 0;	Mismatches 272;	Indels 0;	Gaps 0;
QY	ATTGCCAAGATGTTACGGAATTGATGGTAAAAACCCATATTGATATCTAATPAACCT	170			
DB	ATTCGCCAAGACGTCACCCAGCTCATCGGCAACACCAATGAGTATCTCAACAACCTC	300			
QY	GCGAATGTTGTGTGCGCCGGGTTCGCTPAACTGAGTGTGATGAGGCCATGCTAGT	230			
DB	GTCAAGGATCTGTGCGCAATGTGCTGTAGCTCGAATTAATGAGACCTGCTGTAGC	360			
QY	GTGAAGACAGGATTTGGGTTATGATGATTTGCTGATCAGAGAGAAAGGACTTAATCA	290			
DB	GTGAAGACAGGATTAAGGTTACATATATTAATGATCTGTAAACGAAGGGCTTATTAAT	420			
QY	CTGGAAGAGGTGCTCTCATTTGAGCCCAACAGTGTATATCTGGCATTTGATTAAGCTTC	350			
DB	CTGGAAGAGGTGTTTGTGTGGAAGCAACAGTGAAGAAACAGAGCATTTGGCTTGCTTC	480			
QY	ATGSCACACGACAGGGGTTACAAAGCTCATATTCAATGCTGCTGTATATGAGCTTAG	410			
DB	ATTCTGCTTTTAAGATATTAAGCTATTAAGCTATTAACATATGCTTATCAATGAGCATGAG	540			
QY	AGAAATTCATTCTATTAGCTTTTGGAGCTGAGTGTGTTGTAACATCTCTGTAAAGGA	470			
DB	AGAGAGTCTCTCTCAAGCTTTTGTGTGTAACCTTGTCTTATCTGATGCTGCAAAAGGG	600			
QY	ATGAAAGTCTGTTGGAAGGCTGAGAGATATGTGCTAAGAGCCCAATGCTTAATATA	530			
DB	ATGAAAGGCGCTTAGTATTAAGCTACAGATTTTAAACAGACACCAATTTCTAATG	660			
QY	ATATGAAAGGCTCCGATGGGAATAATGATGCAATTTGTTCTGGGATAGGACATGAGTGT	650			
DB	ATCTGGAGATTCAAAGGGAAGGTGATATATTCATTGTTGATGAAATTTGGAACGGGGGG	780			
QY	ACATATACAGGTGTCGAAAAATATCTTAAAGACAGAAATCGAATATATPAAGCTGATTTGT	710			
DB	ACATATATCTGTGTCGCGCCGCTTTCTCAAGAGAGAAAAATCTGGAAATTAAGGTATTTGT	840			
QY	GTGGAACCAAGTGTGAATATCGAGTCTTCACAGAGAAAGCTGCTGCCACAGATTCAA	770			
DB	ATTGAGCTTCTGTAAATPAACATATCTCTCGGTGGAACCTGTGTCAATPAATCATG	900			
QY	GGGATTTGCTGTGTTTATCTCCGTGATCTTGAAGTCAATCTTCTGTGATGATGTT	830			
DB	GGATCGGCGCAGGATTTGTTTCAAGGAATTTGATGATGATTTCTTGTATGAAATAT	960			
QY	CAATATACAGTATGAAAGCAATAGAACTGCAAGTCTTTCGCTTAAGAAGGCTTA	890			
DB	GAGATATCAAGTATGAAAGCTGTTGAGACAGCAAAACAGTTGCTGTTTCAGGAAGATTA	1020			
QY	TTTGTGGAAATATCTTCGAGCTGACGCTGCTGCTTTTCAGATTGCAAAAAGACCA	950			
DB	CTGTGTGGATCTCTCTGTGGGACAGACGCGCGCCGCAATAAAGGTGTGCAAAAGACCA	1080			
QY	GAATAATCCGGGAAGCTTATTTGTTCGCTTTTCCAGCTCGGGGAGAGGTACTGTTC	1010			
DB	GAGATATCTGGAAAGCTGATAGTGTGTGTTCGAGCTCGGCGAGAGGTACTTTCA	1140			

Qy 1011 TCCGTCCTATTGAGTCAGTCAGACGCGAAGCTGAAGCAATGACTTTGAGCCCTGA 1067  
Db 1141 TCTGTCTCTATCATGTCATTAAGAGAAATGCGAGAACATGCGAGCCCTGAGCCCATGA 1197

Search completed: May 21, 2004, 22:15:45  
Job time : 707 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      May 18, 2004, 13:00:59 ; Search time 60 Seconds
              (without alignments)
              1530.466 Million cell updates/sec
```

Title: US-09-931-457A-31

Perfect score: 1623  
Sequence: 1 MAVESGIAKDYBILGKTP.....LSVLFESEVREAESEMTFEP 3229

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%

```
Database :
1: genesetp1980s:*
2: genesetp1990s:*
3: genesetp2000s:*
4: genesetp2001s:*
5: genesetp2002s:*
6: genesetp2003s:*
7: genesetp2003s:*
8: genesetp2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	1623	100.0	325	2	AAW81018	AAW81018	Fragment
	2	1431	88.2	325	2	AAK49830	AAK49830	Cysteine
	3	1384	85.3	322	5	AAQ29607	AAQ29607	Arabidops
	4	1384	85.3	322	5	ABE94000	ABE94000	Herbicid
	5	1384	85.3	322	5	AAE14774	AAE14774	Arabidops
	6	1379	85.0	322	5	ABE94010	ABE94010	Herbicid
	7	1379	85.0	322	5	AAE14775	AAE14775	Arabidops
	8	1368	84.3	325	7	ABM73727	ABM73727	DNA clone
	9	1368	84.3	418	7	ABM73998	ABM73998	DNA clone
	10	1296	79.9	305	5	ABE93998	ABE93998	Herbicid
	11	1271	78.3	383	2	AAK63756	AAK63756	Sprnach c
	12	1261	77.7	392	5	AAQ09478	AAQ09478	Arabidops
	13	1255	77.3	442	5	ABE94007	ABE94007	Herbicid
	14	1252	77.1	315	5	ABE94013	ABE94013	Herbicid
	15	1252	77.1	315	5	ABE94001	ABE94001	Herbicid
	16	1241	76.5	285	3	AAQ29608	AAQ29608	Arabidops
	17	1239	76.3	324	5	AAQ29609	AAQ29609	Arabidops
	18	1229	75.7	324	5	ABE94016	ABE94016	Herbicid
	19	1229	75.7	324	5	ABE94002	ABE94002	Herbicid
	20	1229	75.7	324	5	AAE14776	AAE14776	Arabidops
	21	1215	74.9	305	3	AAQ09474	AAQ09474	Arabidops
	22	1196	73.7	323	3	AAQ25438	AAQ25438	Arabidops
	23	1192.5	73.5	424	5	ABE94012	ABE94012	Herbicid
	24	1192.5	73.5	424	5	ABE93997	ABE93997	Herbicid
	25	1192	73.4	323	3	AAQ26543	AAQ26543	Arabidops

45	1135	69.19	304	3	Ag06649	Arabidops
44	1140	70.2	304	3	Ag44512	Arabidops
43	1142	70.4	304	3	Ag38511	Arabidops
42	1142	70.4	304	3	Ag26544	Arabidops
41	1142	70.4	284	3	Ag09480	Arabidops
40	1145	70.5	304	3	Ag32539	Arabidops
39	1149	70.8	323	3	Ag36615	Arabidops
38	1159	72.2	392	5	AgB94006	Herbicida
37	1172	72.6	399	5	AgB94003	Herbicida
36	1179	72.6	399	3	Ag35891	Arabidops
35	1179	72.6	324	7	AdC6418	Arabidops
34	1179	72.6	324	5	AgE14779	Arabidops
33	1179	72.6	324	5	AgE14779	Arabidops
32	1179	72.6	324	3	Ag35892	Arabidops
31	1184	73.0	324	3	Ag06648	Arabidops
30	1191	73.4	333	3	Ag44510	Arabidops
29	1191	73.4	323	3	Ag44511	Arabidops
28	1192	73.4	332	3	Ag38509	Arabidops
27	1192	73.4	323	5	AgB94014	Herbicida
26	1192	73.4	323	3	Ag38510	Arabidops
25	1192	73.4	323	3	Ag38510	Arabidops
24	1192	73.4	323	3	Ag38509	Arabidops
23	1192	73.4	323	3	Ag38509	Arabidops
22	1192	73.4	323	3	Ag38509	Arabidops
21	1192	73.4	323	3	Ag38509	Arabidops
20	1192	73.4	323	3	Ag38509	Arabidops
19	1192	73.4	323	3	Ag38509	Arabidops
18	1192	73.4	323	3	Ag38509	Arabidops
17	1192	73.4	323	3	Ag38509	Arabidops
16	1192	73.4	323	3	Ag38509	Arabidops
15	1192	73.4	323	3	Ag38509	Arabidops
14	1192	73.4	323	3	Ag38509	Arabidops
13	1192	73.4	323	3	Ag38509	Arabidops
12	1192	73.4	323	3	Ag38509	Arabidops
11	1192	73.4	323	3	Ag38509	Arabidops
10	1192	73.4	323	3	Ag38509	Arabidops
9	1192	73.4	323	3	Ag38509	Arabidops
8	1192	73.4	323	3	Ag38509	Arabidops
7	1192	73.4	323	3	Ag38509	Arabidops
6	1192	73.4	323	3	Ag38509	Arabidops
5	1192	73.4	323	3	Ag38509	Arabidops
4	1192	73.4	323	3	Ag38509	Arabidops
3	1192	73.4	323	3	Ag38509	Arabidops
2	1192	73.4	323	3	Ag38509	Arabidops
1	1192	73.4	323	3	Ag38509	Arabidops

## ALIGNMENTS

RESULT 1  
AAW61018

AC AAHB1018;

DT 27-SEP-1999 (first entry)

Fragment of cysteine synthase from soybean

Biosynthesis; biosynthetic pathway; lysine; threonine; methionine; amino acid; homoserine kinase.

KW aspartic semialdehyde dehydrogenase; diamnopyrimelate decarboxylase;  
crystallin synthase; glutathione beta-lyase: gene expression: screening

inhibition.

OS Glycine max.  
XY

PN W09856935-A2  
XX

PD 17-DEC-1998

11-JUN-1998  
PR  
XX

PR 12-JUN-1997

AA  
PA  
(DUPO ) DU E

AA  
PI      Falco SC, 1

DR WFI; 1999-08

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
--	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PT encode aspar

PT beta-lyase.

PS Claim 16; Pa

CC Organisation

expression of

CC amino acids

CC production a

of the pathway leading to plant biosynthesis of lysine, threonine, cysteine and isoleucine suggests that over-reduction of expression of genes encoding enzymes involved in the biosynthetic pathway could be used to alter the level of these amino acids in human food and animal feed. This may increase the availability of human food and animal feed by increasing the accumulation of specific free amino acids. The enzymes

CC include aspartic semialdehyde dehydrogenase, homoserine kinase,  
 CC diaminopimelate decarboxylase, cysteine synthase and cystathionine beta-  
 CC lyase. The nucleic acids encoding these enzymes can be used for altering  
 CC the level of expression of the enzymes and for evaluating compounds for  
 CC their ability to inhibit the enzymes' activity  
 XX

SO Sequence 325 AA:

Query Match 100.0%; Score 1623; DB 2; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVERSGIADVTLEIKTFLVYLNKADGCVARVAANKLEMEPCSSVKDRIGYSMTADA 60  
 DB 1 MAVERSGIADVTLEIKTFLVYLNKADGCVARVAANKLEMEPCSSVKDRIGYSMTADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSIERIILAFGAELY 120  
 DB 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSIERIILAFGAELY 120  
 QY 121 LTPDPAKMGKAVOKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIKGSDGKIDAFV 180  
 DB 121 LTPDPAKMGKAVOKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIKGSDGKIDAFV 180  
 QY 181 SGIGTGTTTGAGKYLKEQNPNIKLGVEPVSPVLSGKPGPHKIQIGAGFIPGVLEV 240  
 DB 181 SGIGTGTTTGAGKYLKEQNPNIKLGVEPVSPVLSGKPGPHKIQIGAGFIPGVLEV 240  
 QY 241 NLLDEVQISSDEALIEETAKLALKEGLFVGISSGAAAAPQIARPENAGKLIYAVFPS 300  
 DB 241 NLLDEVQISSDEALIEETAKLALKEGLFVGISSGAAAAPQIARPENAGKLIYAVFPS 300  
 QY 301 FGERYLSVLPFESVREARASMTPEP 325  
 DB 301 FGERYLSVLPFESVREARASMTPEP 325

RESULT 2  
 AAR49830  
 ID AAR49830 standard; protein; 325 AA.

XX AAR49830;  
 AC  
 XX  
 DT 12-OCT-1994 (first entry)  
 XX  
 XX Cysteine synthase.  
 DB Cysteine synthase; plant; expression; probe.  
 XX  
 KM Cysteine synthase; plant; expression; probe.  
 XX  
 OS Spinacia oleracea.  
 XX  
 XX JF06038770-A.  
 XX  
 PD 15-FEB-1994.  
 XX  
 PF 05-FEB-1992; 92JP-00020315.  
 XX  
 PR 05-FEB-1992; 92JP-00020315.  
 XX  
 XX (MITS) MITSUBISHI CORP.  
 PA (MITU) MITSUBISHI KASEI CORP.  
 XX  
 DR MPI: 1994-094834/12.  
 XX  
 DR N-PSDB; AAQ44450.

PT Novel gene coding cysteine synthase - used to increase the cysteine  
 PT content of an agricultural plant.  
 XX  
 PS Claim 2; Page 4-6; 6pp; Japanese.  
 XX  
 CC The cysteine content in an agricultural product can be increased by  
 CC expressing the cysteine synthase in a plant. Probes V822 and V812 used in  
 CC the isolation of the gene are given in AAQ44483-84

XX  
 SO Sequence 325 AA:

Query Match 88.2%; Score 1431; DB 2; Length 325;  
 Best Local Similarity 86.7%; Pred. No. 6.6e-131;  
 Matches 281; Conservative 24; Mismatches 19; Indels 0; Gaps 0;

QY 1 MAVERSGIADVTLEIKTFLVYLNKADGCVARVAANKLEMEPCSSVKDRIGYSMTADA 60  
 DB 1 MBEKRFIAKDVTEILGKTPVLYNTVADGCVARVAANKLEMEPCSSVKDRIGYSMTADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSIERIILAFGAELY 120  
 DB 61 EKSGLITPESVLIETSGNTGIGLAFMAAARGYKLIITMPASMSIERIILAFGAELY 120  
 QY 121 LTPDPAKMGKAVOKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIKGSDGKIDAFV 180  
 DB 121 LTPDPAKMGKAVOKAEIILAKTPNAVYILQGFENPANKVHYETTGPEIKGSDGKIDAFV 180  
 QY 181 SGIGTGTTTGAGKYLKEQNPNIKLGVEPVSPVLSGKPGPHKIQIGAGFIPGVLEV 240  
 DB 181 SGIGTGTTTGAGKYLKEQNPNIKLGVEPVSPVLSGKPGPHKIQIGAGFIPGVLEV 240  
 QY 241 NLLDEVQISSDEALIEETAKLALKEGLFVGISSGAAAAPQIARPENAGKLIYAVFPS 300  
 DB 241 NLLDEVQISSDEALIEETAKLALKEGLFVGISSGAAAAPQIARPENAGKLIYAVFPS 300  
 QY 301 FGERYLSVLPFESVREARASMTPEP 324  
 DB 301 FGERYLSVLPFESVREARASMTPEP 324

RESULT 3  
 AAG29607  
 ID AAG29607 standard; protein; 322 AA.

XX AAG29607;  
 AC  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 35257.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 KM termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 XX  
 PR 05-MAR-1999; 99US-0123180P.  
 XX  
 PR 09-MAR-1999; 99US-0123548P.  
 XX  
 PR 23-MAR-1999; 99US-0125788P.  
 XX  
 PR 25-MAR-1999; 99US-0126264P.  
 XX  
 PR 29-MAR-1999; 99US-0126785P.  
 XX  
 PR 01-APR-1999; 99US-0127462P.  
 XX  
 PR 06-APR-1999; 99US-0128234P.  
 XX  
 PR 08-APR-1999; 99US-0128714P.  
 XX  
 PR 16-APR-1999; 99US-0129845P.  
 XX  
 PR 19-APR-1999; 99US-0130077P.  
 XX  
 PR 21-APR-1999; 99US-0130449P.  
 XX  
 PR 23-APR-1999; 99US-0130510P.  
 XX  
 PR 28-APR-1999; 99US-0131495P.  
 XX  
 PR 30-APR-1999; 99US-0132048P.  
 XX  
 PR 04-MAY-1999; 99US-0132407P.  
 XX  
 PR 05-MAY-1999; 99US-0132485P.

Mon May 24 08:18:53 2004

us-09-931-457a-31.rag

Page 3

PR 06-MAY-1999; 99US-012486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135623P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139493P.  
PR 18-JUN-1999; 99US-0139484P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145226P.  
PR 27-JUL-1999; 99US-0145291P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147312P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147483P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149920P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 13-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160757P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161559P.  
 PR 26-OCT-1999; 99US-0161560P.  
 PR 26-OCT-1999; 99US-0161561P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 85.3%; Score 1384; DB 3; Length 322;  
 Best Local Similarity 83.4%; Pred. No. 2.5e-126;  
 Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;

QY 6 SGIAKVTELTIGTPTLVYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAERKGL 65  
 DB 3 SRIADVTETLIGTPTLVYLNVAEGCVARVAAKLELMEPCSSVKDRIGYSMTADAERKGL 62  
 QY 66 ITPGKSVLEPTSGNTGIGLAFMAARGYLLITPASMSELRRIITLAFGAEVLVTDPA 125  
 DB 63 IKRGESVLEPTSGNTGIGLAFMAARGYLLITPASMSELRRIITLAFGAEVLVTDPA 122  
 QY 126 KMKGAQVQABEILAKTPNAYILQFENPANKVETTTGPEITWKSODGKIDAFVSGIGT 185  
 DB 123 KMKGAQVQABEILAKTPNAYILQFENPANKVETTTGPEITWKSODGKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYLKEONPNKILGVEPVESPVLSGKPGPHKIOGICAGFIPGVLVNLDE 245  
 DB 183 GGTITGAGKYLKEONPNKILGVEPVESPVLSGKPGPHKIOGICAGFIPGVLVNLDE 242  
 QY 246 VVOISSDEAIEFTAKLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIYAVFPGERY 305  
 DB 243 VVOISSDESIDMAROLALKEGLVGISSGAAAAAIKLAOREPBNAGKLFVALFPGERY 302  
 QY 306 LSSVLPESVREARESMTEF 324  
 DB 303 LSTVLPDTRKERABAMTFE 321

RESULT 4  
 ABB94000  
 ID ABB94000 standard; protein; 322 AA.  
 AC ABB94000;  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Herbicidally active polypeptide SEQ ID NO 3211.  
 XX  
 KW Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN MO200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001MO-EP009892.  
 XX  
 PR 28-AUG-2001; 2001MO-EP009892.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Tietjen K, Weidner M;  
 XX  
 DR WPI; 2002-269010/31.  
 XX

PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 PS Claim 5; SEQ ID NO 3211; 261pp + Sequence Listing; English.  
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides

Sequence 322 AA;  
 Query Match 85.3%; Score 1384; DB 5; Length 322;  
 Best Local Similarity 83.4%; Pred. No. 2.5e-126;  
 Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;

QY 6 SGIAKVTELTIGTPTLVYLNKLDGCVARVAAKLELMEPCSSVKDRIGYSMTADAERKGL 65  
 DB 3 SRIADVTETLIGTPTLVYLNVAEGCVARVAAKLELMEPCSSVKDRIGYSMTADAERKGL 62  
 QY 66 ITPGKSVLEPTSGNTGIGLAFMAARGYLLITPASMSELRRIITLAFGAEVLVTDPA 125  
 DB 63 IKRGESVLEPTSGNTGIGLAFMAARGYLLITPASMSELRRIITLAFGAEVLVTDPA 122  
 QY 126 KMKGAQVQABEILAKTPNAYILQFENPANKVETTTGPEITWKSODGKIDAFVSGIGT 185  
 DB 123 KMKGAQVQABEILAKTPNAYILQFENPANKVETTTGPEITWKSODGKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYLKEONPNKILGVEPVESPVLSGKPGPHKIOGICAGFIPGVLVNLDE 245  
 DB 183 GGTITGAGKYLKEONPNKILGVEPVESPVLSGKPGPHKIOGICAGFIPGVLVNLDE 242  
 QY 246 VVOISSDEAIEFTAKLALKEGLFVGISSGAAAAAFOIAKRPENAGKLIYAVFPGERY 305  
 DB 243 VVOISSDESIDMAROLALKEGLVGISSGAAAAAIKLAOREPBNAGKLFVALFPGERY 302  
 QY 306 LSSVLPESVREARESMTEF 324  
 DB 303 LSTVLPDTRKERABAMTFE 321

RESULT 5  
 AAE14774  
 ID AAE14774 standard; protein; 322 AA.  
 AC AAE14774;  
 DT 30-OCT-2002 (first entry)  
 XX  
 DE Arabidopsis thaliana cysteine synthase #1.  
 XX  
 KW Herbicide; cysteine synthase; plant growth; developmental abnormality;  
 XX  
 EC 4.2.99.8.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN MO200246451-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 25-OCT-2001; 2001MO-US050890.  
 XX  
 PR 26-OCT-2000; 2000US-00697225.  
 XX  
 PA (PARA-) PARADIGM GENETICS INC.  
 XX



PI Klotz A, Woessner U, Zayed A, Boyes D, Davis K, Hamilton C;  
PI Ascenzi R, Hoffman N;  
DR WPI: 2002-519673/55.

XX Identifying compound as candidate herbicide; comprises contacting  
PT cysteine synthase with compound and detecting presence and/or absence of  
PT binding between compound and synthase.

PS Claim 4; Page 25-26; 30pp; English.

XX The invention relates to a method of identifying a compound as a  
CC candidate for a herbicide, comprising contacting cysteine synthase with  
CC the compound, and detecting the presence and/or absence of binding  
CC between the compound and cysteine synthase, where binding indicates that  
CC the compound is a candidate for a herbicide. Cysteine synthase is  
CC essential for plant growth. Specifically, the inhibition of cysteine  
CC synthase gene expression in plant seedlings results in severe  
CC developmental abnormalities. Thus, cysteine synthase is used as a target  
CC for identification of herbicides. The present sequence is Arabidopsis  
CC thaliana cysteine synthase (EC 4.2.99.8) Atocacy

XX Sequence 322 AA;

Query Match 85.3%; Score 1384; DB 5; Length 322;  
Best Local Similarity 83.4%; Pred. No. 2.5e-126;  
Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;

QY 6 SGIAQVTELTGKPTPLVYLNKADGCVARVAAKLEIMEPCSSVDRIGYSMIADAEEKGL 65  
DB 3 SRINADVTELTGNTPLVYLNNAAGCGVAAKLEIMEPCSSVDRIGYSMIADAEEKGL 62  
QY 66 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITPASMISLERRIILAFGAELVLTDP 125  
DB 63 IKPGESVLIIEPTSGNTGIGLAFMAAARGYKLIITPASMISLERRIILAFGAELVLTDP 122  
QY 126 KMGKAVOKAEELIAKTPNAYILQOFENPANKYHETTGPEIKMGSDGKIDAFVSGIGT 185  
DB 123 KMGKALINAAEELIAKTPNGYMLQOFENPANKYHETTGPEIKMGSDGKIDGFSIGIGT 182  
QY 186 GGTITGAGKYLKEONPNIKLIGVEPVESPVLSGKPGPHKIQGIAGGFIPELVNLTDE 245  
DB 183 GGTITGAGKYLKEONPNVNLIGVEPVESPVLSGKPGPHKIQGIAGGFIPELVNLTDE 242  
QY 246 VVOISSDEALITAKLAKELFVGISSGAAAFAAFOIAKRPENAGKLIIVAFSPFGERY 305  
DB 243 VVOVSSDESIDIMARQALAKELGLVGISSGAAAFAAFOIAKRPENAGKLIIVAFSPFGERY 302  
QY 306 LSSVLFESVRRRAESMTFE 324  
DB 303 LSTVLPDATRKEAAMTFE 321

RESULT 6  
AB94010 ID ABB94010 standard; protein; 322 AA.

XX ABB94010;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 3221.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX

PR 28-AUG-2001; 2001WO-EP009892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidner M;

DR WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms.

PS Claim 5; SEQ ID NO 3221; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (AB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides

XX Sequence 322 AA;

Query Match 85.0%; Score 1379; DB 5; Length 322;  
Best Local Similarity 83.1%; Pred. No. 7.8e-126;  
Matches 265; Conservative 30; Mismatches 24; Indels 0; Gaps 0;

QY 6 SGIAQVTELTGKPTPLVYLNKADGCVARVAAKLEIMEPCSSVDRIGYSMIADAEEKGL 65  
DB 3 SRINADVTELTGNTPLVYLNNAAGCGVAAKLEIMEPCSSVDRIGYSMIADAEEKGL 62  
QY 66 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITPASMISLERRIILAFGAELVLTDP 125  
DB 63 IKPGESVLIIEPTSGNTGIGLAFMAAARGYKLIITPASMISLERRIILAFGAELVLTDP 122  
QY 126 KMGKAVOKAEELIAKTPNAYILQOFENPANKYHETTGPEIKMGSDGKIDAFVSGIGT 185  
DB 123 KMGKALINAAEELIAKTPNGYMLQOFENPANKYHETTGPEIKMGSDGKIDGFSIGIGT 182  
QY 186 GGTITGAGKYLKEONPNIKLIGVEPVESPVLSGKPGPHKIQGIAGGFIPELVNLTDE 245  
DB 183 GGTITGAGKYLKEONPNVNLIGVEPVESPVLSGKPGPHKIQGIAGGFIPELVNLTDE 242  
QY 246 VVOISSDEALITAKLAKELFVGISSGAAAFAAFOIAKRPENAGKLIIVAFSPFGERY 305  
DB 243 VVOVSSDESIDIMARQALAKELGLVGISSGAAAFAAFOIAKRPENAGKLIIVAFSPFGERY 302  
QY 306 LSSVLFESVRRRAESMTFE 324  
DB 303 LSTVLPDATRKEAAMTFE 321

RESULT 7  
AAE14775 ID AAE14775 standard; protein; 322 AA.

XX AAE14775;

XX 30-OCT-2002 (first entry)

XX Arabidopsis thaliana cysteine synthase #2.

XX Herbicide; cysteine synthase; plant growth; developmental abnormality;  
XX EC 4.2.99.8.

XX Arabidopsis thaliana.

XX WO200246451-A2.

XX



```

XX  ABM73998;
AC
XX  17-OCT-2003 (first entry)
DT
XX  DNA clone originating in barley containing SNP sequence #408.
DE
XX  Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX  Hordeum vulgare.
XX  WO2003057877-A1.
XX  17-JUL-2003.
XX  16-DEC-2002; 2002WO-1B005403.
XX  20-DEC-2001; 2001JP-00387059.
XX  20-DEC-2001; 2001JP-00387131.
XX  20-DEC-2001; 2001JP-00403299.
XX  20-DEC-2001; 2001JP-00403300.
XX  27-SEP-2002; 2002JP-00327515.
XX  (UYN1-) UNIV JAPAN OKAYAMA.
XX  Sato K, Takeda K, Kohara Y;
PI
XX  WPI; 2003-587127/55.
XX  Single nucleotide polymorphism sites in barley varieties and DNA
PT  sequences containing them for analysis and identification of barley
PT  varieties and production of barley transformants with desired
PT  characteristics.
XX  Disclosure; SEQ ID XX; 284bp; Japanese.
XX  The present invention relates to oligonucleotide clones originating in
CC  barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC  (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC  varieties, identification of particular varieties and genotype-phenotype
CC  analysis, isolation of specific genes and creation of new varieties by
CC  transformation of barley varieties with them and production of new barley
CC  varieties with desired properties. The present sequence represents an
CC  oligonucleotide clone sequence featured in the specification. The
CC  sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published-pct-sequences
XX  SQ
XX  Sequence 418 AA;
XX  Query Match      84.3%; Score 1368; DB 7; Length 418;
XX  Best Local Similarity 83.0%; Pred. No. 1,4e-124;
XX  Matches 263; Conservative 26; Mismatches 28; Indels 0; Gaps 0;

```

```

QY  308 SVLPESVREAESMTFF 324
DB  402 SVLPSTIKKESMTVVR 418
XX  RESULT 10
XX  ID ABB93998 standard; protein; 305 AA.
XX  ABB93998;
XX  31-MAY-2002 (first entry)
XX  31-MAY-2002 (first entry)
XX  Herbicidally active polypeptide SEQ ID NO 3209.
XX  Herbicidally active polypeptide SEQ ID NO 3209.
XX  Herbicidal; plant; agriculture; herbicide.
XX  Arabidopsis thaliana.
XX  WO200210210-A2.
XX  07-FEB-2002.
XX  28-AUG-2001; 2001WO-EP009892.
XX  28-AUG-2001; 2001WO-EP009892.
XX  (FARB ) BAYER AG.
XX  Tiejien K, Weidner M;
XX  WPI; 2002-269010/31.
XX  Identifying plant target proteins for herbicidally active compounds,
PT  comprising aligning and comparing nucleic acid or amino acid sequences
PT  from plant with nucleic acid or amino acid sequences from non-plant
PT  organisms.
XX  Claim 5; SEQ ID NO 3209; 261pp + Sequence Listing; English.
XX  The invention relates to identifying target proteins (ABB90790-ABB94016)
CC  for herbicidally active compounds, comprising aligning and comparing
CC  nucleic acid or amino acid sequences from plant with nucleic acid or
CC  amino acid sequences from non-plant organisms using suitable search
CC  parameters, where plant sequences having an E-value greater by a factor
CC  of 3 than the E-value of most similar non-plant sequences are selected.
CC  The polypeptides or nucleic acids encoding them are useful for
CC  identifying modulators. The identified modulators are useful as
CC  herbicides
XX  SQ
XX  Sequence 305 AA;
XX  Query Match      79.9%; Score 1296; DB 5; Length 305;
XX  Best Local Similarity 78.2%; Pred. No. 9.2e-118;
XX  Matches 254; Conservative 28; Mismatches 23; Indels 20; Gaps 1;

```

QY 241 NLDEVOQISSDEALETKALKEGLFVGISGAAAAAFOIAKRPENACKLIIVAFPS 300  
 Db 221 DLIDEVQVSSSEESIDNARILAREGLVIGISSGAAATAAIKAKRPENACKLIIVAFPS 280  
 QY 301 FGERYLSVLPESVREAEASMTFEP 325  
 Db 281 FGERYLSVLPFDARKREAEITMTFEP 305

## RESULT 11

AA063756  
 ID AAR63756 standard; protein; 383 AA.

AA063756;

DT 13-JUN-1995 (first entry)

DE Spinach cysteine synthase.

KM Cysteine; spinach; spinach oleracea; pUC19; M13mpl8; plant; nutrition;  
 KM feed value; pharmaceutical; food additive; cosmetic; E.coli.

OS Spinacia oleracea.

PN JP06245773-A.

PD 06-SEP-1994.

PR 26-FEB-1993; 93JP-00038527.

PR 26-FEB-1993; 93JP-00038527.

XX (MITS) MITSUBISHI CORP.

PA (MITU) MITSUBISHI KASEI CORP.

DR WPI; 1994-321282/40.

DR N-PSDB; AA074413.

PS Claim 2; Page 4-6; 6pp; Japanese.

CC The amino acid sequence of the novel cysteine synthase from

CC spinach, Spinacia oleracea. The probes (AA074414-5) were used to obtain a

CC clone from a cDNA library derived from spinach seedling leaves RNA. The

CC 1.5 kb insert was ligated into the cloning vectors pUC19 and M13 mp18.

CC The gene, 1463 bp, encodes a protein of 383 a.a. The cysteine synthase

CC gene is expressed in plants to elevate the cysteine content in the plant.

CC The nutritional and feed value of the plant are expected to be enriched.

CC Cysteine, one of the S-containing amino acids, can be used as materials

CC for various pharmaceuticals, food additives or cosmetics. Production of

CC the protein is a useful step in the production of cysteine

XX

XX Sequence 383 AA:

Query March 78.3%; Score 1271; DB 2; Length 383;

Best Local Similarity 73.4%; Pred. No. 3.6e-115; Mismatches 33; Indels 0; Gaps 0;

Matches 234; Conservative 52; Mismatches 33; Indels 0; Gaps 0;

QY 3 VERSGIADVTELLIGKTPLYVINTLADQVAVAKLEMEPCSSVDRIIGYSMLDAEE 62

Db 62 IEGNIAEDVSQILGKTPMYVINTLVNYSKGSVANIKALESMEPCSSVDRIIGYSMLDAEQ 121

QY 63 KGIITPKSVLIEPTSGNTGIGLAFAAAGYKLIITPMSLSERITIIILAGSALVLT 122

Db 122 KGVITPKSVLIEPTSGNTGIGLAFAAAGYKLIITPMSLSERITIIILAGSALVLT 161

QY 123 DPAGMGAVOKAEELIAKTPNAVILLOFENPANPKVYEITGPEIKGSDGKIDAFVSG 182

Db 182 DPAGMGAVOKAEELIAKTPDSYMLQOFDPANPKIHYETTGPEIWEITKGVKIDFVAG 241

QY 183 IGTGTTTGAKTKAKGNPNIKLIGVFPVPSVLSGKRPENKIQIGAGPIFGVLEVN 242

Db 242 IGTGTTTGAKTKAKGNPNIKLIGVFPVPSVLSGKRPENKIQIGAGPIFGVLEVN 301  
 QY 243 LDEVOQISSDEALETKALKEGLFVGISGAAAAAFOIAKRPENACKLIIVAFPSFG 302  
 Db 302 MDVIRVSSSEAEVEMAKQIAKREGLVIGISSGAAAAAARIGKRPENACKLIIVAFPSFG 361  
 QY 303 ERTYLSVLPESVREAEASMTFEP 321  
 Db 362 ERTYLSVLPESVREAEASMTFEP 380

## RESULT 12

AA09478  
 ID AAG09478 standard; protein; 392 AA.

AA09478;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 7430.

KM Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

KM termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PR 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 15-APR-1999; 99US-0129645P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130445P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 04-MAY-1999; 99US-0132048P.

PR 06-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134570P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135333P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141877P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142302P.
PR	09-JUL-1999;	99US-0142803P.
PR	12-JUL-1999;	99US-0142877P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147433P.
PR	09-AUG-1999;	99US-0148171P.
PR	09-AUG-1999;	99US-0148172P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148555P.
PR	13-AUG-1999;	99US-0148584P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149829P.
PR	23-AUG-1999;	99US-0149802P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154038P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155148P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156566P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157351P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158237P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159648P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
PR	09-AUG-1999;	99US-0147433P.
PR	09-AUG-1999;	99US-0148171P.
PR	09-AUG-1999;	99US-0148172P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148555P.
PR	13-AUG-1999;	99US-0148584P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149829P.
PR	23-AUG-1999;	99US-0149802P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154038P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155148P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156566P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157351P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158237P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159648P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query March 77.7%; Score 1261; DB 3; Length 392;  
 Best Local Similarity 74.9%; Pred. No. 3.5e-114;  
 Matches 239; Conservative 41; Mismatches 39; Indels 0; Gaps 0;

QY 3 VERSGIADVTETLIGKTELVYLNKLDGCVAAKLEMEPCSSVDRIGYSMTIADAE 62  
 DB 70 VEGNTIADNAQILIGKTPMYLNNVKGCVASVAALKEIMEPCSSVDRIGYSMTIADAE 129  
 QY 63 KGLTTPKSVLLEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGAEVL 122  
 DB 130 KGLTTPKSVLLEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAFGAEVL 189  
 QY 123 DPAGKMGAVQKAEELIATKTNAYILQFENPAMPKHYETTGPEIMKSGDKIDAIVSG 182  
 DB 190 BPAGKMGAVQKAEELIATKTNAYILQFENPAMPKHYETTGPEIMKSGDKIDAIVSG 249  
 QY 183 IGTGCTTGAGKTYLKEONPNIKLIGVBPVSVLSGKPGPHKIQIGAGFIPGVLEVN 242  
 DB 250 IGTGCTTGAGKTYLKEONPNIKLIGVBPVSVLSGKPGPHKIQIGAGFIPGVLEVN 309  
 QY 243 IDEVVOISDRAITFAKLLALKEGLFVGISGAAAAAFAQAKPENAGKELIVAPPSFG 302  
 DB 310 VDEITAISSERAIETSKQALQEGLVGISSGAAAAAIVAKPENAGKELIVAPPSFG 369  
 QY 303 ERYLSSVLFESVREASMS 321  
 DB 370 ERYLSTOLFQSIREECEOM 388

## RESULT 13

ABB94007  
 ID ABB94007 standard; protein; 442 AA.  
 AC ABB94007;  
 XX  
 DT 31-MAY-2002 (first entry)  
 DE Herbicidally active polypeptide SEQ ID NO 3218.  
 XX  
 KM Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W0200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PR 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Tietjen K, Weidler M;  
 XX  
 DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 PS Claim 5; SEQ ID NO 3218; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 CC  
 CC SQ Sequence 442 AA;

Query Match 77.3%; Score 1255; DB 5; Length 442;

Best Local Similarity 71.3%; Pred. No. 1.6e-113;  
 Matches 241; Conservative 43; Mismatches 36; Indels 18; Gaps 2;

QY 2 AYERS-----GIKAVTETLIGKTELVYLNKLDGCVAAKLEMEPCSSVDRIGYS 55  
 DB 101 AYRETEGPDGLNADNVSQLIGKTPMYLNNVKGCVASVAALKEIMEPCSSVDRIGYS 160  
 QY 56 MIDAEKGLITPGKSVLLEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAF 115  
 DB 161 MYTDAQKQKISGKSVLLEPTSGNTGIGLAFMAAARGYKLIITMPASMSLEERILLAF 220  
 QY 116 GAELVITDPKMGKGA VQKAEELIATKTNAYILQFENPAMPKHYETTGPEIMKSGDK 175  
 DB 221 GAELVITDPKMGKGA VQKAEELIATKTNAYILQFENPAMPKHYETTGPEIMKSGDK 280  
 QY 176 IDAFVSGIGTGITTAGVYLNKLDGCVAAKLEMEPCSSVDRIGYSMTIADAE 235  
 DB 281 VDLFVAGIGTGITTAGVYLNKLDGCVAAKLEMEPCSSVDRIGYSMTIADAE 340  
 QY 236 GVLAVNLDEVVOISDRAITFAKLLALKEGLFVGISGAAAAAFAQAKPENAGKELIV 295  
 DB 341 KMLDQKIMDEVVOISDRAITFAKLLALKEGLFVGISGAAAAAFAQAKPENAGKELIV 400  
 QY 296 -----AVPSPGERTYLSVLFESVREASMS 321  
 DB 401 VTVNSMNAVFOVFPSPGERTYLSVLFESVREASMS 438

## RESULT 14

ABB94013  
 ID ABB94013 standard; protein; 315 AA.  
 AC ABB94013;  
 XX  
 DT 31-MAY-2002 (first entry)  
 DE Herbicidally active polypeptide SEQ ID NO 3224.  
 XX  
 KM Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W0200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PR 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Tietjen K, Weidler M;  
 XX  
 DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 PS Claim 5; SEQ ID NO 3224; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 CC  
 CC XX

SQ Sequence 315 AA;

Query Match 77.1%; Score 1252; DB 5; Length 315;

Best Local Similarity 80.3%; Pred. No. 1.9e-113; Mismatches 245; Conservative 28; Indels 2; Gaps 1;

QY 6 SGIADVTLLIGTLPVYLNKLDGCVARVAARVAAKLEMEPCSSVKDRIGYSMTADAEEKL 65  
 DB 3 SRIADVTLLIGTLPVYLNKLDGCVARVAARVAAKLEMEPCSSVKDRIGYSMTADAEEKL 62  
 QY 66 ITPGKSVLIPTSGNTGIGLAFAAARGVKLIITTPASMSLERRITLLAFGAELVLTDP 125  
 DB 63 IKPGSVLIPTSGNTGIGLAFAAARGVKLIITTPASMSLERRITLLAFGAELVLTDP 122  
 QY 126 KGMKGAIVQAEIILAKTNPAYIILOFPENPANKHYETTTGPEIMKSGDGIKIDAFVSGIGT 185  
 DB 123 KGMKGAIVQAEIILAKTNPAYIILOFPENPANKHYETTTGPEIMKSGDGIKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYLKEONPNIKLIGVEPVESPVLSGKRGPHKIQIGAGFIPGVLEVNLD 245  
 DB 183 G--ITGAGSILKNRANVKLYGVEPVESALILIGKRGPHKIQIGAGFIPGVLEVNLD 240  
 QY 246 VVOISSDEAIEFTAKLLALKEGLFVGISGAAAAAFAQIAKRPENNAKLIIVAVPSPFGERY 305  
 DB 241 VVOVSSDESIDMARQALKEGLFVGISGAAAAAFAQIAKRPENNAKLIIVAVPSPFGERY 300  
 QY 306 LSSVL 310  
 DB 301 LSTVL 305

## RESULT 15

ID ABB94001 standard; protein, 315 AA.

AC ABB94001;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3212.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN MO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidner M;

DR WPI, 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
 comprising aligning and comparing nucleic acid or amino acid sequences  
 from plant with nucleic acid or amino acid sequences from non-plant  
 organisms.

XX Claim 5; SEQ ID NO 3212; 261pp + Sequence listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 for herbicidally active compounds, comprising aligning and comparing  
 nucleic acid or amino acid sequences from plant with nucleic acid or  
 amino acid sequences from non-plant organisms using suitable search  
 parameters, where plant sequences having an E-value greater by a factor  
 of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for

CC Identifying modulators. The identified modulators are useful as  
 herbicides

SQ Sequence 315 AA;

Query Match 77.1%; Score 1252; DB 5; Length 315;

Best Local Similarity 80.3%; Pred. No. 1.9e-113; Mismatches 245; Conservative 28; Indels 2; Gaps 1;

QY 6 SGIADVTLLIGTLPVYLNKLDGCVARVAARVAAKLEMEPCSSVKDRIGYSMTADAEEKL 65  
 DB 3 SRIADVTLLIGTLPVYLNKLDGCVARVAARVAAKLEMEPCSSVKDRIGYSMTADAEEKL 62  
 QY 66 ITPGKSVLIPTSGNTGIGLAFAAARGVKLIITTPASMSLERRITLLAFGAELVLTDP 125  
 DB 63 IKPGSVLIPTSGNTGIGLAFAAARGVKLIITTPASMSLERRITLLAFGAELVLTDP 122  
 QY 126 KGMKGAIVQAEIILAKTNPAYIILOFPENPANKHYETTTGPEIMKSGDGIKIDAFVSGIGT 185  
 DB 123 KGMKGAIVQAEIILAKTNPAYIILOFPENPANKHYETTTGPEIMKSGDGIKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYLKEONPNIKLIGVEPVESPVLSGKRGPHKIQIGAGFIPGVLEVNLD 245  
 DB 183 G--ITGAGSILKNRANVKLYGVEPVESALILIGKRGPHKIQIGAGFIPGVLEVNLD 240  
 QY 246 VVOISSDEAIEFTAKLLALKEGLFVGISGAAAAAFAQIAKRPENNAKLIIVAVPSPFGERY 305  
 DB 241 VVOVSSDESIDMARQALKEGLFVGISGAAAAAFAQIAKRPENNAKLIIVAVPSPFGERY 300  
 QY 306 LSSVL 310  
 DB 301 LSTVL 305

Search completed: May 18, 2004, 13:07:06  
 Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:04:09 ; Search time 21 Seconds  
(without alignments)  
1488.678 Million cell updates/sec

Title: US-09-931-457A-31

Sequence: 1 MAVERSGIKDVTTELICKTP.....LSSVLFEVRRREASMTFEP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 beqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

	Minimum	Match	0%
Post-processing:			

```

Database :
PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

A>Title: Molecular cloning of a cysteine synthase cDNA from Citrullus vulgaris (watermelon)
A>Reference number: S46438; MUID:94316193; PMID:8041362
A>Accession: S46438
A>Molecule type: mRNA
A>Residues: 1-325 <NOJ>

```

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1453	89.5	325	1	566438	cysteine synthase
2	1435	88.4	325	2	T07001	cysteine synthase
3	1431	88.2	325	2	S35094	cysteine synthase
4	1384	85.3	322	2	A71412	cysteine synthase
5	1378	84.9	325	1	SS2738	cysteine synthase
6	1364	84.0	325	1	US0762	cysteine synthase
7	1314	81.0	386	2	T07002	cysteine synthase
8	1271	78.3	387	2	T52650	cysteine synthase
9	1270	78.3	352	2	T07962	probable cysteine synthase
10	1261	77.7	383	2	S29373	cysteine synthase
11	1261	77.7	392	2	A64870	cysteine synthase
12	1255	77.3	442	2	T47800	cysteine synthase
13	1252	77.1	315	2	S48654	cysteine synthase
14	1249	77.0	374	1	A43407	cysteine synthase
15	1229	75.7	324	2	S49586	cysteine synthase
16	1179	72.6	324	2	T52609	cysteine synthase
17	1170	72.1	390	2	T09000	cysteine synthase
18	1163	71.7	392	2	S48655	cysteine synthase
19	1047	64.5	368	2	A55450	cysteine synthase
20	1017	62.7	368	2	T47936	cysteine synthase
21	1004	61.9	319	2	AB2121	cysteine synthase
22	989	60.9	320	2	AB2374	cysteine synthase
23	971	59.8	331	2	S77347	cysteine synthase
24	968	59.6	344	2	T19367	cysteine synthase
25	949.5	58.5	310	2	G70660	cysteine synthase
26	947	58.3	322	2	AH2614	cysteine synthase
27	947	58.3	322	2	G97396	cysteine synthase
28	940.5	57.9	330	2	T44912	cysteine synthase
29	925	57.0	317	2	T23591	cysteine synthase

30	916	56.4	336	2	AH3264	cysteine synthase
31	912	56.2	330	2	H81161	cysteine synthase
32	896.5	55.2	309	2	E97175	cysteine synthase
33	882	54.3	322	2	C82258	cysteine synthase
34	878	54.1	337	2	C89009	cysteine synthase
35	877	54.0	308	2	T44614	cysteine synthase
36	869.5	53.6	329	1	S55321	cysteine synthase
37	865.5	53.3	333	2	AD0810	cysteine synthase
38	862	53.1	322	2	A10363	cysteine synthase
39	860.5	53.0	328	2	G87698	cysteine synthase
40	859	52.9	308	2	S66103	cysteine synthase
41	857.5	52.8	333	1	SYCBAC	cysteine synthase
42	857.5	52.8	333	2	A85884	cysteine synthase
43	857.5	52.8	333	2	F91039	cysteine synthase
44	856	52.7	334	2	B83306	cysteine synthase
45	851.5	52.5	333	1	SYB2AC	cysteine synthase

## ALIGNMENTS

## RESULT 1

cysteine synthase (EC 4.2.99.8) - watermelon

C:\Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

R;Noji, M.; Murakoshi, I.; Saito, K.

**A;Title:** Molecular cloning of a c

A/Accession: S46438

A;Residues: 1-325 <NC

A;Note: the source is designated as *Citrus* *vulgaris*

C<sub>3</sub>Superfamily: threonine dehydratase

P;49/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Protein	Residue	Binding site	Pyridoxal phosphate (Lys)	(covalent)	#status predicted
P49/Binding site	pyridoxal phosphate (Lys)	(covalent)			

Query Match	89.5%	Score 1453	DB 1	Length 325
Best Local Similarity	87.1%	Pred. No. 1.6e-99		
Matches 283	Conservative 24	Mismatches 18	Indels 0	Gaps 0
QY	1	MAVERSGIADVTIELKTEPLVYLNKLADGCVAAVAAKLELMERCSYKDKRIGYSMTIDA	60	
DB	1	MAAKRSIADVDVIELGNTPLVYLNRAVDDGVAAVAAKLELMERCSYKDKRIGYSMTIDA	60	
QY	61	BEKGLIIPGSKSVLIEPTSGNTGIGLAFMAAARGKYLITTPASMSLEKRIILLAFGAELY	120	
DB	61	ENGLIIPGSSVLIHPTSGNTGIGLAFIAAKGKRLIICMPASMSLEKRIILLAFGAELY	120	
QY	121	LTPAPKMKKAVQKAEIILAKTPHAYLILQGFENANRKHVETTGPELTKRSDGKIDIAFY	180	
DB	121	LTPAPKMKKAVQKAEIILAKTPHAYLILQGFENANRKHVETTGPELTKRSDGKIDIAFY	180	
QY	181	SGIGTGTITGAGKTYLAKBNPNIKLIGVEPEVSEFVLSGCKRGPFKIKIGAGFIPGVLEV	240	
DB	181	SGIGTGTITGAGKTYLAKBNPNIKLIGVEPEVSEFVLSGCKRGPFKIKIGAGFIPGVLEV	240	
QY	241	NILDEVAUUSDSDAIRPAKLLALKEGIFVGISGAAAMAAAFQIAKRPENNGKLIIVAVPFS	300	
DB	241	NILDEVAUUSDSDAIRPAKLLALKEGILVGISGAAAMAAAFQIAKRPENNGKLIIVAVPFS	300	
QY	301	FGERYLSSVLFESVRRRAESMTPEP	325	
DB	301	FGERYLSTVLFESVKKETEMVPEP	325	



C/Species: Solanum tuberosum (potato)  
 C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Jun-1999  
 C/Accession: T07001  
 R/Hesse, H.  
 Submitted to the EMBL Data Library, January 1998  
 A/Reference number: Z13854  
 A/Accession: T07001  
 A/Status: translated from GB/EMBL/DDBT  
 A/Molecule type: mRNA  
 A/Residues: 1-325 <S&S>  
 A/Cross-references: EMBL:AF044172; NID:G3290019; PIDN:AAC25635.1; PID:G3290020  
 A/Experimental source: cv. Beroлина  
 A/Function:  
 A/Description: forms cysteine from O-acetyl-serine and hydrogen sulfide and releases co  
 A/Pathway: cysteine biosynthesis  
 C/Superfamily: threonine dehydratase  
 C/Keywords: carbon-oxygen lyase; cytosol; phosphoprotein; pyridoxal phosphate  
 F/45/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 88.4%; Score 1435; DB 2; Length 325;  
 Best Local Similarity 85.2%; Pred. No. 3.4e-98;  
 Matches 277; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTADA 60  
 DB 1 MAGEKIGIAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTIDA 60  
 QY 61 EEKGLITPKGSVLIETPSGNTGIGLAPMAAGYKLIITPMASMSLERRIILAPAEIV 120  
 DB 61 EEKGLITPKGSVLIETPSGNTGIGLAPMAAGYKLIITPMASMSLERRIILAPAEIV 120  
 QY 121 LTDPKAMKGAVOAKAEILAKTPNAYILIQFENPANKVHYETTGPEIKWGSDDKIDAV 180  
 DB 121 LTDPKAMKGAVOAKAEILAKTPNAYILIQFENPANKVHYETTGPEIKWGSDDKIDAV 180  
 QY 121 SGITGCTTTGAGKYLKEQNPVNIKLIQVEPVESPVLSGGKPPHKIQGIGAGTIPGVLEV 240  
 DB 121 SGITGCTTTGAGKYLKEQNPVNIKLIQVEPVESPVLSGGKPPHKIQGIGAGTIPGVLEV 240  
 QY 181 NLDEVOISSDEAIEETAKLALKEGLFVGISGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 DB 181 NLDEVOISSDEAIEETAKLALKEGLFVGISGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 QY 241 NLIDVVOISSSESIEMAKLALKEGLLVGISSGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 DB 241 NLIDVVOISSSESIEMAKLALKEGLLVGISSGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 QY 301 FGERYISSVLPESVRRAESMTPE 325  
 DB 301 FGERYISSVLPESVRRAESMTPE 325

## RESULT 3

S35094  
 Cysteine synthase (EC 4.2.99.8) A - spinach  
 N/Alternate names: O-acetylserine (thiol)-lyase  
 C/Species: Spinacia oleracea (spinach)  
 C/Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Jun-2000  
 A/Accession: S35094; S35095  
 R/Saito, K.; Mizur, N.; Yamazaki, M.; Hirano, H.; Murakoshi, I.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 8078-8082, 1992  
 A/Title: Molecular cloning and bacterial expression of cDNA encoding a plant cysteine sy  
 A/Reference number: S35094; MUID:92390392; PMID:11518833  
 A/Accession: S35094  
 A/Molecule type: mRNA  
 A/Residues: 1-325 <S&S>  
 A/Cross-references: EMBL:DL0476; NID:G218276; PIDN:BA01279.1; PID:G218277  
 A/Experimental source: strain cv. pardee; tissue leaf  
 A/Accession: S35095  
 A/Molecule type: protein  
 A/Residues: 5-11,71-90,163-182,246-262 <S&S>  
 C/Function:  
 A/Pathway: cysteine biosynthesis  
 C/Superfamily: threonine dehydratase  
 C/Keywords: carbon-oxygen lyase; phosphoprotein; pyridoxal phosphate  
 F/225/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 88.2%; Score 1431; DB 2; Length 325;  
 Best Local Similarity 86.7%; Pred. No. 6.7e-98;  
 Matches 281; Conservative 24; Mismatches 19; Indels 0; Gaps 0;

QY 1 MAVERSGIAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTADA 60  
 DB 1 MVEKAFIAKDVTLLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTIDA 60  
 QY 61 EEKGLITPKGSVLIETPSGNTGIGLAPMAAGYKLIITPMASMSLERRIILAPAEIV 120  
 DB 61 EESGLITPKGSVLIETPSGNTGIGLAPMAAGYKLIITPMASMSLERRIILAPAEIV 120  
 QY 121 LTDPKAMKGAVOAKAEILAKTPNAYILIQFENPANKVHYETTGPEIKWGSDDKIDAV 180  
 DB 121 LTDPKAMKGAVOAKAEILAKTPNAYILIQFENPANKVHYETTGPEIKWGSDDKIDAV 180  
 QY 181 NLDEVOISSDEAIEETAKLALKEGLFVGISGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 DB 181 NLDEVOISSDEAIEETAKLALKEGLFVGISGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 QY 241 NLIDVVOISSSESIEMAKLALKEGLLVGISSGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 DB 241 NLIDVVOISSSESIEMAKLALKEGLLVGISSGMAAAAFQIAKPEENAGKLIIVAVPS 300  
 QY 301 FGERYISSVLPESVRRAESMTPE 324  
 DB 301 FGERYISSVLPESVRRAESMTPE 324

## RESULT 4

A71412  
 Cysteine synthase (EC 4.2.99.8) 3A, cytosolic - Arabidopsis thaliana  
 N/Alternate names: cytosolic O-acetylserine (thiol) lyase; O-acetylserine sulfhydrylas  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 A/Variety: columbia  
 C/Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
 A/Reference number: A71412; S65533  
 R/Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Di  
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Wellenreger, T.; Pohl, T.M.; Terry, N.; Gi  
 Nave, T.; Hempel, S.; Kotter, P.; Entian, R.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
 Nature 391, 485-488, 1998  
 A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Putdomenek  
 erhof, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; A  
 C.; Chalatzis, N.  
 A/Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th  
 A/Reference number: A71400; MUID:98121113; PMID:9461215  
 A/Accession: A71412  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-322 <REV>  
 A/Cross-references: GB:297337; NID:G2244829; PIDN:CA010267.1; PID:G2244845  
 R/Bertoso, C.; Vega, U.K.; Goto, C.  
 FEBS Lett. 363, 1-5, 1995  
 A/Title: A new member of the cytosolic O-acetylserine(thiol)lyase gene family in Arabid  
 A/Reference number: S65533; MUID:95246850; PMID:7729527  
 A/Accession: S65533  
 A/Molecule type: mRNA  
 A/Residues: 1-272, 'E', 274-322 <BAR>  
 A/Cross-references: EMBL:X84097; NID:G804949; PIDN:CA568893.1; PID:G804950  
 C/Genetics:  
 A/Gene: cys-3A  
 A/Map position: 4COP9-4G3845  
 C/Superfamily: threonine dehydratase  
 C/Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phos  
 F/46/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 85.3%; Score 1384; DB 2; Length 322;  
 Best Local Similarity 83.4%; Pred. No. 1.9e-94;  
 Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;

QY 6 SGIAKDVTELLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTADAEEKL 65  
 DB 3 SRIADVTELLIGTPTLVYLNKLDGCVARVAAKLEMEPCSSVDRIGYSMTADAEEKL 62

QY 66 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 125  
 DB 63 ITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 122  
 QY 126 KMKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 165  
 DB 123 KMKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 245  
 DB 183 GGTITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 242  
 QY 246 VVOISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 305  
 DB 243 VVOISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 302  
 QY 306 LSSVLFESVRRBAESMTPE 324  
 DB 303 LSTVLPDAIRKBAESMTPE 321

RESULT 5  
 S52738  
 Cysteine synthase (EC 4.2.99.8) precursor - maize  
 C1Species: Zea mays (maize)  
 C1Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C1Accession: S52738  
 R.Brander, K.A.; Owttrim, G.W.; Brumold, C.  
 submitted to the EMBL Data Library, March 1995  
 A:Description: Isolation of a putative plastidic isoform of cysteine synthase from maize  
 A:Reference number: S52738  
 A:Accession: S52738  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-325 <DB>  
 A:Cross-references: EMBL:X85803; NID:G758352; PIDD:CAA59798.1; PID:G758353  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; phosphoprotein; pyridoxal phosphate  
 F19/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 84.9%; Score 1378; DB 1; Length 325;

Best Local Similarity 84.3%; Pred. No. 5.3e-94;

Matches 268; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

QY 8 IAKDVTIELIGTPIVLYLANKADGCVARVAKLELMEPCSSVYKDRIGYSMIDAEKGLIT 67  
 DB 8 IAKDVTIELIGTPIVLYLANKADGCVARVAKLELMEPCSSVYKDRIGYSMIDAEKGLIT 67  
 QY 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 127  
 DB 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 127  
 QY 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 127  
 DB 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 127  
 QY 128 MKANQKAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 167  
 DB 128 MKANQKAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 167  
 QY 128 MKANQKAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 167  
 DB 128 MKANQKAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 167  
 QY 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 247  
 DB 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 247  
 QY 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 247  
 DB 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 247  
 QY 248 QISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 307  
 DB 248 QISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 307  
 QY 308 SVLFESVRRBAESMTPE 325  
 DB 308 SVLFESVRRBAESMTPE 325

RESULT 6  
 JS0762  
 Cysteine synthase (EC 4.2.99.8) precursor - wheat

N:Alternate names: O-acetylserine (thiol)-lyase  
 C1Species: Triticum aestivum (common wheat)  
 C1Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
 C1Accession: JS0762  
 R.Youssefian, S.; Nakamura, M.; Sano, H.  
 submitted to JIPID, September 1992  
 A:Reference number: JS0762  
 A:Accession: JS0762  
 A:Molecule type: mRNA  
 A:Residues: 1-325 <DB>  
 A:Cross-references: DDBJ:D13153; NID:G218334; PIDD:BA02438.1; PID:G218335  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; plastid; pyridic  
 F11-37/Domain: transit peptide (plastid) #status predicted <TMP>  
 F138-325/Product: cysteine synthase #status predicted <MAT>  
 F1226/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 84.0%; Score 1364; DB 1; Length 325;

Best Local Similarity 82.3%; Pred. No. 5.7e-93;

Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKDVTIELIGTPIVLYLANKADGCVARVAKLELMEPCSSVYKDRIGYSMIDAEKGLIT 67  
 DB 9 IAKDVTIELIGTPIVLYLANKADGCVARVAKLELMEPCSSVYKDRIGYSMIDAEKGLIT 68  
 QY 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 127  
 DB 68 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 128  
 QY 69 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 128  
 DB 69 PKGSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAELVLTDBA 128  
 QY 128 MKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 187  
 DB 128 MKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 188  
 QY 129 MKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 188  
 DB 129 MKGAQVQAEETLAKTAPNAVILQEPENPANKVHYETTGPRIMKSGDCKIDAFVSGIGT 188  
 QY 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 247  
 DB 188 TITGAGKYLKEQNPNIKLIGVEPVESPVLSGSGKPEPHKIQIGAGFIPGVLEVNLTDE 248  
 QY 248 QISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 307  
 DB 248 QISSDEALETAKLALKEGLFVGISSGAAAQIAQIKRPNNAKLIIVANPSPGERTY 308  
 QY 308 SVLFESVRRBAESMTPE 324  
 DB 308 SVLFESVRRBAESMTPE 325

RESULT 7

Cysteine synthase (EC 4.2.99.8) precursor, chloroplast - potato

C1Species: Solanum tuberosum (potato)

C1Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Jun-1999

C1Accession: T07002

submitted to the EMBL Data Library, January 1998

A:Reference number: Z15854

A:Accession: T07002

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-386 <HBS>

A:Cross-references: EMBL:AF044173; NID:G3290021; PIDD:AAC25636.1; PID:G3290022

A:Experimental source: cv. Beiroina

A:Genome: nuclear

A:Function: forms cysteine from O-acetyl-serine and hydrogen sulfide and releases CO<sub>2</sub>

C:Superfamily: threonine dehydratase

C:Keywords: carbon-oxygen lyase; chloroplast; cysteine biosynthesis; phosphoprotein; py

F110/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 81.0%; Score 1314; DB 2; Length 386;

Best Local Similarity 78.1%; Pred. No. 3.4e-89;

Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;

QY 3 VERSGIADVTIELIGTPLYINKLADGCVARVAAKLEIMPCSSVKDRIGYSMIADAE 62  
 DB 64 IEENIAEDVTOLIGTNPVYNTIAGCVANIAAKLEIMPCSSVKDRIGYSMIADAE 123  
 QY 63 KGLITPCKSLIPTSNTGICLAFMAAARGYLLITMPASMSLERIITLAAAGALVLT 122  
 DB 124 KGLISPEKTLVETSTNTGICLAFIAPASRGYLLITMPASMSLERIITLAAAGALVLT 183  
 QY 123 DPAKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDKIDAFVSG 182  
 DB 184 DPAKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDKIDAFVSG 243  
 QY 183 IGTGGITTGAKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 242  
 DB 244 IGTGGITTGAKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 303  
 QY 243 LDEVVOISSDEAIFETAKLALKEGLFVGISGGAAMAAAFQIAXRPENAGLIVAFVPSFG 302  
 DB 304 MDEVIEISSDEAIFETAKLALKEGLFVGISGGAAMAAAFQIAXRPENAGLIVAFVPSFG 363  
 QY 303 ERYLSSVLFESVREAESEW 321  
 DB 364 ERYLSSVLFESVREAESEW 382

## RESULT 8

TS2650  
 Cysteine synthase (EC 4.2.99.8) precursor, mitochondrion [validated] - Arabidopsis thaliana  
 N:Alternate names: O-acetylserine (thiol) lyase  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 03-Nov-2000  
 C/Accession: T52650  
 R:Jost, R., Berkowitz, O., Wirtz, M., Hopkins, L., Hawkesford, M.J., Hell, R.  
 Gene 253, 237-247, 2000  
 A>Title: Genomic and functional characterization of the *cas* gene family encoding O-acetylserine synthase  
 A/Reference number: Z26157  
 A/Accession: T52650  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-387 <TOS>  
 A/Cross-references: EMBL:AF271727, PIDD:CA871290.1  
 A:Experimental source: cultivar Columbia  
 A:Gene: *cas*  
 A:Genome: oase  
 A:Function:  
 A:Description: EC 4.2.99.8 [validated, MVID:20400348]; forms cysteine from O-acetyl-serine  
 A:Pathway: cysteine biosynthesis  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; mitochondrion  
 F:1-36/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
 F:37-387/Product: cysteine synthase #status predicted <MAT>

Query Match 78.3%; Score 1271; DB 2; Length 387;  
 Best Local Similarity 73.9%; Pred. No. 5e-86; Indels 6; Gaps 1;  
 Matches 24; Conservative 43; Mismatches 36;

QY 2 AVERS-----GIADVTIELIGTPLYINKLADGCVARVAAKLEIMPCSSVKDRIGYS 55  
 DB 58 AVRETEPDGIANADNVSQILIGTNPVYNTIAGCVANIAAKLEIMPCSSVKDRIGYS 117  
 QY 56 MINDABEKGLITGKSLIPTSNTGICLAFMAAARGYLLITMPASMSLERIITLAAAGALVLT 115  
 DB 118 MINDABEKGLITGKSLIPTSNTGICLAFIAPASRGYLLITMPASMSLERIITLAAAGALVLT 177  
 QY 116 GAEIVLTDPAKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDK 175  
 DB 178 GAEIVLTDPAKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDK 237  
 QY 176 IDAFVSGITGKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 235  
 DB 238 VDIIVAGITGKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 297

QY 236 GAEIVLTDPAKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDK 295  
 DB 298 KMLDQKIMBIVLALISSEBIEFAKOLALKEGLMVGISGGAAMAAAFQIAXRPENAGLIV 357  
 QY 296 AVFPSSGGERYLSVLFESVREAESEW 321  
 DB 358 AVFPSSGGERYLSVLFESVREAESEW 383

## RESULT 9

T07962  
 probable cysteine synthase (EC 4.2.99.8) 1A precursor - Chlamydomonas reinhardtii  
 N:Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulphydrylase  
 C:Species: Chlamydomonas reinhardtii  
 C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 21-Jan-2000  
 C/Accession: T07962  
 R:Ravina, C.G.; Barroso, C.; Vega, J.M.; Gotor, C.  
 Submitted to the EMBL Data Library, July 1998  
 A:Description: Cysteine biosynthesis in chlamydomonas reinhardtii. Molecular cloning a  
 A/Reference number: Z16250  
 A/Accession: T07962  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-352 <RAV>  
 A/Cross-references: EMBL:AF078693, NID:G3342568, PIDD:AACT7794.1; PID:G3342569  
 C:Genetics:  
 A:Gene: *cys-1A*  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phos  
 F:1-31/Domain: transit peptide (chloroplast) #status predicted <TNP>  
 F:32-352/Product: cysteine synthase 1A #status predicted <MAT>  
 F:74/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 78.3%; Score 1270; DB 2; Length 352;  
 Best Local Similarity 74.4%; Pred. No. 5.2e-86; Indels 0; Gaps 0;  
 Matches 24; Conservative 41; Mismatches 42;

QY 2 AVERS-----GIADVTIELIGTPLYINKLADGCVARVAAKLEIMPCSSVKDRIGYSMIADAE 61  
 DB 27 AAVKNIADVTIELIGTNPVYNTIAGCVANIAAKLEIMPCSSVKDRIGYSMIADAE 86  
 QY 62 EKGITPCKSLIPTSNTGICLAFMAAARGYLLITMPASMSLERIITLAAAGALVLT 121  
 DB 87 EKGITPCKSLIPTSNTGICLAFIAPASRGYLLITMPASMSLERIITLAAAGALVLT 146  
 QY 122 TDPKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDKIDAFVSG 181  
 DB 147 TDPKMGAVOKABEILAKTPNAYILQOFNPANPKVHYETTGPEIWKSGDKIDAFVSG 206  
 QY 182 GIGGTTTGAKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 241  
 DB 207 GIGGTTTGAKTYLKEONPNIKLIGVFPVSPVLSGKRGPHKIQIGAGFTIGVLEVL 266  
 QY 242 LDEVVOISSDEAIFETAKLALKEGLFVGISGGAAMAAAFQIAXRPENAGLIVAFVPSFG 301  
 DB 267 LDEVVOISSDEAIFETAKLALKEGLFVGISGGAAMAAAFQIAXRPENAGLIVAFVPSFG 326  
 QY 302 GERYLSSVLFESVREAESEW 325  
 DB 327 GERYLSSVLFESVREAESEW 380

## RESULT 10

S29733  
 Cysteine synthase (EC 4.2.99.8) B precursor, chloroplast - spinach  
 N:Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulphydrylase  
 C:Species: Spinacia oleracea (spinach)  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000  
 C/Accession: S29733; S33591; S62218; S23944  
 R:Rolland, N.; Droux, M.; Lebun, M.; Douce, R.  
 Arch Biochem Biophys 300, 213-222, 1993  
 A>Title: O-acetylserine(thiol) lyase from spinach (Spinacia oleracea L.) leaf: cDNA clon  
 A/Reference number: S29733; MVID:93143317; PMID:8424655



Query Match 77.3%; Score 1255; DB 2; Length 442;  
 Best Local Similarity 71.3%; Pred. No. 8.9e-85;  
 Matches 241; Conservative 43; Mismatches 36; Indels 18; Gaps 2;

QY 2 AVES-----GIKADVTETLIGKTPLYVINKLADGCYARVAAKLEMEPCSSVDRIGYS 55  
 DB 101 AVKSTFGDGLINDVNSQIGKTPMYLINSIAAGCVANIAKLEMEPCSSVDRIGYS 160  
 QY 56 MIADAEKGLITPKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAF 115  
 DB 161 MYTDAEGGPIISPGKSVLIEPTSGNTGIGLAFIAASGYRLITPASMISERRIILAF 220  
 QY 116 GAELVLDPAKMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDK 175  
 DB 221 GAELVLDPAKMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDK 280  
 QY 176 IDAFVSGIGTGTITGAGKYLKEQNPNIKIGVPEVPSVPLSGKPGPHKIQIGAGFIP 235  
 DB 281 VDIIVAGIGTGTITGAGKYLKEQNPNIKIGVPEVPSVPLSGKPGPHKIQIGAGFIP 340  
 QY 236 GVLVNLIDEVQIISDEALETAKLAKKGLIFGISSGAAAAAFAQIARPENAGKLI 295  
 DB 341 KMLDQKIMDEVIALISSEALHTAKQLKKEGLMWGISSGAAAAAIVAKRPENAGKLI 400  
 QY 296 -----AVPSPGERRYLSVLFESVREAREASM 321  
 DB 401 VTVNEMANAVQVVPFSPGRYLTSPFQSIREEVEOM 438

## RESULT 13

48694  
 cysteine synthase (EC 4.2.99.8) isoform 5-8, cytosolic - Arabidopsis thaliana  
 N:Alternate names: O-acetylserine (thiol) lyase  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 16-Feb-1995 #sequence\_revision 13-Jan-1996 #text\_change 22-Jun-1999  
 C/Accession: S48694  
 R:Hell, R.; Bork, C.; Bogdanova, N.; Frolov, I.; Hauschild, R.  
 FEBS Lett. 351, 257-262, 1994  
 A:Title: Isolation and characterization of two cDNAs encoding for compartment specific A  
 A:Reference number: S48694; MUID:94364481; PMID:8082776  
 A:Accession: S48694  
 A:Molecule type: mRNA  
 A:Residues: 1-315 <HEU>  
 A:Cross-references: EMBL:X80376; NID:9603229; PID:CAAS6553.1; PID:9560130  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; cytosol; homodimer; phosphoprote  
 F:46/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.1%; Score 1252; DB 2; Length 315;  
 Best Local Similarity 80.3%; Pred. No. 9.5e-85;  
 Matches 245; Conservative 28; Mismatches 30; Indels 2; Gaps 1;  
 QY 6 SGIAKDVTEIGKTPLYVINKLADGCYARVAAKLEMEPCSSVDRIGYSMTDAERKGL 65  
 DB 3 SRIKDVTEIGKTPLYVINKLADGCYARVAAKLEMEPCSSVDRIGYSMTDAERKGL 62  
 QY 66 ITTGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLDPA 125  
 DB 63 IKPESVLIIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLDPA 122  
 QY 126 KMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDGKIDAFVSGIGT 185  
 DB 123 KMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDGKIDAFVSGIGT 182  
 QY 186 GGTGAGKYLKEQNPNIKIGVPEVPSVPLSGKPGPHKIQIGAGFIPGVLVNLID 245  
 DB 183 G-ITGAGSILKNTANVYKLVGVPESVPLISGKPGPHKIQIGAGFIPSVNLVLD 240  
 QY 246 VVOISSEALHTAKLAKKGLIFGISSGAAAAAFAQIARPENAGKLIIVAFSPGERY 305  
 DB 241 VVOISSEALHTAKLAKKGLIFGISSGAAAAAFAQIARPENAGKLIIVAFSPGERY 300

QY 306 LSSVL 310  
 DB 301 LSTVL 305

## RESULT 14

A43407  
 cysteine synthase (EC 4.2.99.8) precursor - pepper  
 N:Alternate names: O-acetylserine (thiol) lyase; O-acetylserine sulfhydrylase  
 C:Species: Capsicum annuum (pepper)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C/Accession: A43407; S24637  
 R:Romer, S.; d'Harlingue, A.; Camara, B.; Schantz, R.; Kuntz, M.  
 J. Biol. Chem. 267, 17966-17970, 1992  
 A:Title: Cysteine synthase from Capsicum annuum chromoplasts. Characterization and cDN  
 A:Reference number: A43407; MUID:92388158; PMID:1381358  
 A:Accession: A43407  
 A:Molecule type: mRNA  
 A:Residues: 1-374 <ROM>  
 A:Cross-references: GB:X64874; NID:917943; PID:CA46086.1; PID:917944  
 A:Note: sequence extracted from NCBI backbone (NCBI:112876)  
 C:Genetics:  
 A:Gene: nuclear  
 C:Superfamily: threonine dehydratase  
 C:Keywords: carbon-oxygen lyase; chloroplast; cysteine biosynthesis; phosphoprotein; P  
 F:108/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 77.0%; Score 1249; DB 1; Length 374;  
 Best Local Similarity 75.2%; Pred. No. 2e-84;  
 Matches 240; Conservative 38; Mismatches 31; Indels 10; Gaps 1;

QY 3 VERSGIKDVTEIGKTPLYVINKLADGCYARVAAKLEMEPCSSVDRIGYSMTDAER 62  
 DB 62 IBSGIAKDVTEIGKTPLYVINKLADGCYARVAAKLEMEPCSSVDRIGYSMTDAER 121  
 QY 63 KGLITGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLT 122  
 DB 122 KGLITGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLT 181  
 QY 123 DPAKMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDGKIDAFVSG 182  
 DB 182 DPAKMGAVOKAEIILAKTPNAYILQOFENPAPKHYETTPGEIMKSDGKIDAFVSG 241  
 QY 183 IGTGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLT 242  
 DB 242 IGTGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMISERRIILAFGAELVLT 291  
 QY 243 LBEVQIISDEALETAKLAKKGLIFGISSGAAAAAFAQIARPENAGKLIIVAFSPFG 302  
 DB 292 LBEVQIISDEALETAKLAKKGLIFGISSGAAAAAFAQIARPENAGKLIIVAFSPFG 351  
 QY 303 ERTLSVLPESVREAREASM 321  
 DB 352 ERTLSVLPESVREAREASM 370

## RESULT 15

S49586  
 cysteine synthase (EC 4.2.99.8) ACS1 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 27-Jan-1995 #sequence\_revision 24-Feb-1995 #text\_change 22-Jun-1999  
 C/Accession: S49586  
 R:Hesse, H.; Altman, T.  
 submitted to the EMBL Data Library, September 1994  
 A:Description: Molecular cloning of a cysteine synthase cDNA from Arabidopsis thaliana  
 A:Reference number: S49586  
 A:Accession: S49586  
 A:Molecule type: mRNA  
 A:Residues: 1-324 <HES>  
 A:Cross-references: EMBL:X81697; NID:9572518; PID:CA57343.1; PID:9572519  
 C:Genetics:  
 A:Gene: ACS1  
 C:Superfamily: threonine dehydratase

C Keywords: carbon-oxygen lyase; cysteine biosynthesis; phosphoprotein; pyridoxal phosph  
E/49/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 75.7%; Score 1229; DB 2; Length 324;

Best Local Similarity 75.9%; Pred. No. 4,8e-83;

Matches 246; Conservative 34; Mismatches 38; Indels 6; Gaps 4;

QY	6	SGIADVTLEIGKTEPLVYINKLA-DGCVARVAKLEIMEPCS-SVKDRIGYSMTADAE	62
DB	3	SRIADVTLEIGNTPELVIANVSLKGCGRVAKLEIMEPCSVSVDKDRIGFSMISDAEK	62
QY	63	KGLITPGKSVLIEPTSGNTG-IGLAFMAARGYKLIITWPAASLERRIILLAFGAEVL	121
DB	63	KGLITQGSVLEIPTSNTGBILLAPYACKGYKLIITWPAASLERRIILLAFGAEVL	122
QY	122	TDPAGMKAVOKAEIILAKTNATILLOFENPANKVHYETTGPEIKGSDKIDAFVS	181
DB	123	TDPAGMKGSYRKAOEIIAKTENGWMLQFENPANKZIHYPPEIKWGKGAKIDGFVS	182
QY	182	GIGTGGITGAGYKLEQNPNIKLGVEFEVESPVLGGKPGPHKIQIGAGFIPGVLEVN	241
DB	183	GIGTGGITGAGYKLEQNPNIKLGVEFEVESAIIISGKPGPTKIQIGAGFIPSVLVND	242
QY	242	LIDEVVOISSDEAIFAKLLAKELFVGISSGAAAAAFOYAKRPENAGKLIYVFPSPF	301
DB	243	LIDEVVOVSSDESIMARQILKEGLVIGISSG--AAAAIKLQRPENAGKLEVAIFPSF	300
QY	302	GERYLSVLFESVRRRABSMTEP	325
DB	301	GRRYLSVLFEDADKERSRMPSRP	324

Search completed: May 18, 2004, 13:09:04  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:02:39 ; Search time 17 Seconds  
(without alignments)  
995.459 Million cell updates/sec

Title: US-09-931-457A-31

Perfect score: 1623  
Sequence: 1 MAVERSGTAKVTELTGTP.....LSVLFSSVRRAASMTREP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1453	89.5	325	1	CYSK CITLA
2	1435	88.4	325	1	CYSK SOLTU
3	1431	88.2	325	1	CYSK SPIOL
4	1417	87.3	322	1	CYK1 BRABU
5	1409	86.8	324	1	CYK2 BRABU
6	1384	85.3	322	1	CYSK ARATH
7	1378	84.9	325	1	CYSK MAIZE
8	1377	84.8	325	1	CYK2 ORYSA
9	1364	84.0	325	1	CYSK WHEAT
10	1356	83.5	321	1	CYK1 ORYSA
11	1314	81.0	386	1	CYSL SOLTU
12	1271	78.3	383	1	CYSL SPIOL
13	1271	78.3	424	1	CYSM ARATH
14	1261	77.7	392	1	CYSL ARATH
15	1249	77.0	374	1	CYSL CAPAN
16	1045	64.4	404	1	CYCN ARATH
17	971	59.8	312	1	CYSK SYNS3
18	949.5	58.5	310	1	CYSK MYCTU
19	940.5	57.9	310	1	CYSG MYCEL
20	869.5	53.6	329	1	SRPG SYNPF
21	865.5	53.3	322	1	CYSK SALTY
22	859	52.9	307	1	CYSK BACSU
23	857.5	52.8	322	1	CYSK BCOUL
24	816.5	50.3	316	1	CYSK HAEIN
25	815	50.2	310	1	CYSK STAM
26	801	49.4	310	1	CYSK STABP
27	769.5	47.4	322	1	CYSK BUCAP
28	765.5	47.2	315	1	CYSK BUCAI
29	737.5	45.4	311	1	CYSM BACSU
30	701	43.2	307	1	CYSK FLASP
31	682	42.0	299	1	CYSM CAMUS
32	611.5	37.7	303	1	CYSM SALTY
33	610.5	37.6	303	1	CYSM BCOUL

34	598.5	36.9	327	1	CYSM AOUAE	O67507 aquifex aeo
35	585.5	36.1	306	1	CYSM HELPY	P56067 helicobacte
36	577.5	35.6	305	1	CYSM HELPU	O22m66 helicobacte
37	550	33.9	550	1	CBS HUMAN	P35520 homo sapien
38	545	33.6	560	1	CBS RAT	P32232 rattus norv
39	544.5	33.5	497	1	CBS DICDI	P46794 dictyosteli
40	541.5	33.4	507	1	CBS YEAST	P32582 saccharomyc
41	483	29.8	323	1	CYSM MYCTU	O10624 mycobacteri
42	463	28.5	393	1	CYSK YEAST	P53206 saccharomyc
43	459	28.3	371	1	CYSK EMENT	P50867 emerticella
44	433	26.7	155	1	CYSK STABA	O59918 staphylococ
45	432	26.6	395	1	CYSK_SCHPO	P87131 schizosacch

## ALIGNMENTS

RESULT 1	ID	CYSK CITLA	STANDARD	PRT	325 AA.
AC	O43317:				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Cysteine synthase (EC 2.5.1.47) (Beta-pyrazolealanine synthase)				
DE	(Beta-PA/CSase) (EC 2.5.1.51) (L-mimosine synthase) (EC 2.5.1.52) (O-acetylserine sulfhydrylase) (O-acetylserine (thiol)-lyase) (CSase)				
DE	(OAS-ly)				
OC	Citrus limon (Watermelon) (Citrus vulgaris).				
OC	Eutryopota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Cucurbitales; Cucurbitaceae; Citrullus.				
OX	NCBI_TaxID=3654;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Seeding;				
RX	MEDLINE=94316193; PubMed=8041362;				
RA	Noji M., Murakoshi I., Saito K.;				
RT	"Molecular cloning of a cysteine synthase cDNA from Citrullus				
RT	vulgaris (watermelon) by genetic complementation in an Escherichia				
RT	coli Cys-auxotroph.";				
RT	Mol. Gen. Genet. 244:57-66(1994).				
RM	[2]				
RP	FUNCTION.				
RX	MEDLINE=94107305; PubMed=8280125;				
RA	Noji M., Murakoshi I., Saito K.;				
RT	"Evidence for identity of beta-pyrazolealanine synthase with cysteine				
RT	synthase in watermelon: formation of beta-pyrazole-alanine by cloned				
RT	cysteine synthase in vitro and in vivo.";				
RT	Biochem. Biophys. Res. Commun. 197:1111-1117(1993).				
RM	[3]				
RP	FUNCTION.				
RX	MEDLINE=97165913; PubMed=9013806;				
RA	Saito K., Kimura N., Ikegami F., Noji M.;				
RT	"Production of plant non-protein amino acids by recombinant enzymes of				
RT	sequential biosynthetic reactions in bacteria.";				
RT	Biochem. Pharm. Bull. 20:47-53(1997).				
CC	-1- FUNCTION: Produces L-cysteine from O-acetyl-L-serine and hydrogen				
CC	sulfide. Can also use pyrazole and 3,4-dihydroxypyridine instead				
CC	of the hydrogen sulfide to produce two plant specific non-				
CC	protein amino acids beta-pyrazolylalanine and L-mimosine.				
CC	-1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +				
CC	acetate.				
CC	-1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + pyrazole = 3-(pyrazol-				
CC	1-yl)-L-alanine + acetate.				
CC	-1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + 3,4-dihydroxypyridine =				
CC	3-(3,4-dihydroxypyridin-1-yl)-L-alanine + acetate.				
CC	-1- COFACTOR: Pyridoxal phosphate.				
CC	-1- PATHWAY: Cysteine biosynthesis.				
CC	-1- SUBUNIT: Homodimer (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).				
CC	-1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-				
CC	-1- synthase family.				

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D28777; BAA05965.1; -  
DR PIR; S46438; S46438.  
DR HSSP; P12674; 10AS.

DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthk.  
DR InterPro; IPR005856; Cys\_synthk.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMs; TIGR01139; CysK; 1.  
DR TIGRFAMs; TIGR01136; CysK; 1.  
DR PROSITE; PS00901; Cys SYNTHASE; 1.  
DR TRANSFERASE; Cysteine biosynthesis; Pyridoxal phosphate.  
DR BINDING; 49  
DR DOMAIN; 275 280 POLY-ALA.

Query March 325 AA; 34342 MW; 4A91B6FAFB5F539 CRC64;  
Best local Similarity 89.5%; Score 1453; DB 1; Length 325;  
Matches 283; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

QY 1 MAVERSGIANDVTELGKTPVLYLNKLADGCVARVAAKLEMBPCSSVKDRIGYSMTADA 60  
DB 1 MAAKSTIAKDVTELGKTPVLYLNKLADGCVARVAAKLEMBPCSSVKDRIGYSMTADA 60  
QY 61 BEKGLITPKSVLIEPTSGNTGIGLAFMAAAGYKLIITPMSLSERRIITLAFGAEIV 120  
DB 61 ENKGLITPKSVLIEPTSGNTGIGLAFMAAAGYKLIITPMSLSERRIITLAFGAEIV 120  
QY 121 LTPDPAKMGAVOKAEIILAKTPNAYILLOQFENPANKVHYETTGPEIWKSGDKIDAFV 180  
DB 121 LTPDPAKMGAVOKAEIILAKTPNAYILLOQFENPANKVHYETTGPEIWKSGDKIDAFV 180  
QY 181 SGITGTTTGAKTYLKEQNPINKLIGVPEPVSPLSGKPGPHKIQIGAGTIPGVLEV 240  
DB 181 SGITGTTTGAKTYLKEQNPINKLIGVPEPVSPLSGKPGPHKIQIGAGTIPGVLEV 240  
QY 241 NLDEVOIISDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRENNAKLIVAFPS 300  
DB 241 NLDEVOIISDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRENNAKLIVAFPS 300  
QY 301 FGERYLSVLPFVSVRRAESMTPEP 325  
DB 301 FGERYLSVLPFVSVRRAESMTPEP 325

## RESULT 2

CYSK\_SOLTU STANDARD; PRT; 325 AA.

AC 081154;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
DS (O-acetylserine (thiol)-lyase) (CSase A) (CS-A) (OAS-TL A).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OC NCBI\_TaxID=4113;  
OX 11

RP SEQUENCE FROM N. A.  
RC STRAIN=cv. Berolina;  
RA Heese H.; Hoelgen R.;  
RT "Isolation of cDNAs encoding cytosolic and plastidic cysteine synthase

RT isoforms from Solanum tuberosum".  
RL (In) Plant Gene Register PCR98-057.  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate.  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF044172; AAC25635.1; -  
DR PIR; T07001; T07001.  
DR HSSP; P12674; 1FCJ.

DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthk.  
DR InterPro; IPR005856; Cys\_synthk.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMs; TIGR01139; CysK; 1.  
DR TIGRFAMs; TIGR01136; CysK; 1.  
DR PROSITE; PS00901; Cys SYNTHASE; 1.  
DR TRANSFERASE; Cysteine biosynthesis; Pyridoxal phosphate.  
DR BINDING; 49  
DR DOMAIN; 275 280 POLY-ALA.

Query March 325 AA; 34341 MW; 86119D046C647BC CRC64;  
Best local Similarity 85.2%; Score 1435; DB 1; Length 325;  
Matches 277; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAVERSGIANDVTELGKTPVLYLNKLADGCVARVAAKLEMBPCSSVKDRIGYSMTADA 60  
DB 1 MAAKSTIAKDVTELGKTPVLYLNKLADGCVARVAAKLEMBPCSSVKDRIGYSMTADA 60  
QY 61 BEKGLITPKSVLIEPTSGNTGIGLAFMAAAGYKLIITPMSLSERRIITLAFGAEIV 120  
DB 61 ENKGLITPKSVLIEPTSGNTGIGLAFMAAAGYKLIITPMSLSERRIITLAFGAEIV 120  
QY 121 LTPDPAKMGAVOKAEIILAKTPNAYILLOQFENPANKVHYETTGPEIWKSGDKIDAFV 180  
DB 121 LTPDPAKMGAVOKAEIILAKTPNAYILLOQFENPANKVHYETTGPEIWKSGDKIDAFV 180  
QY 181 SGITGTTTGAKTYLKEQNPINKLIGVPEPVSPLSGKPGPHKIQIGAGTIPGVLEV 240  
DB 181 SGITGTTTGAKTYLKEQNPINKLIGVPEPVSPLSGKPGPHKIQIGAGTIPGVLEV 240  
QY 241 NLDEVOIISDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRENNAKLIVAFPS 300  
DB 241 NLDEVOIISDEAIFETAKLALKEGLFVGISGAAAAAFOIAKRENNAKLIVAFPS 300  
QY 301 FGERYLSVLPFVSVRRAESMTPEP 325  
DB 301 FGERYLSVLPFVSVRRAESMTPEP 325

## RESULT 3

CYSK\_SPIOL STANDARD; PRT; 325 AA.

AC 000834;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
DS (O-acetylserine (thiol)-lyase) (CSase A) (OAS-TL A).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OC NCBI\_TaxID=4113;  
OX 11



OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Caryophyllales; Amaranthaceae; Spinacia.  
 OK NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 5-11; 71-90; 163-182 AND 246-262.  
 RC STRAIN=cv. Parade; TISSUE=Leaf;  
 RX MEDLINE=92390392; PubMed=1518833;  
 RA Saito K., Miura N., Yamazaki M., Hirano H., Murakoshi I.;  
 RT "Molecular cloning and bacterial expression of cDNA encoding a plant  
 RT cysteine synthase-";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8078-8082(1992).  
 RN [2]  
 RP MUTAGENESIS OF LYSINE RESIDUES.  
 RX MEDLINE=93345669; PubMed=8344414;  
 RA Saito K., Kurosawa M., Murakoshi I.;  
 RT "Determination of a functional lysine residue of a plant cysteine  
 RT synthase by site-directed mutagenesis, and the molecular evolutionary  
 RT implications";  
 RL FEBS Lett. 328:111-114(1993).  
 CC -I- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate;  
 CC -I- COFACTOR: Pyridoxal phosphate.  
 CC -I- PATHWAY: Cysteine biosynthesis.  
 CC -I- SUBUNIT: Homodimer.  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -I- TISSUE SPECIFICITY: Leaves and roots.  
 CC -I- PTM: The N-terminus is blocked.  
 CC -I- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D10476; BAA01279.1; -.  
 DR PIR; S35094; S35094.  
 DR HSSP; P12674; 10KS.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synth.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; cyrK; 1.  
 DR TIGRFAMs; TIGR01136; cyrKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 DR Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.  
 FT BINDING 49  
 FT DOMAIN 275 280  
 FT POLY-ALA.  
 FT SEQUENCE 325 AA; 34186 MW; 161B467B670DEB6 CRC64;  
 Query Match 88.2%; Score 1431; DB 1; Length 325;  
 Best Local Similarity 86.7%; Pred. No. 6; 7e-91;  
 Matches 281; Conservative 24; Mismatches 19; Indels 0; Gaps 0

```

Db      181  SGIGTGCTTGAGAGYKLEQNDPVDTLGLBEVESVIVSGKRPKXIQGLGAGFIPGVLDV 240
Cy      241  NLIDVEVOISDSEKLETFKTLALKEGFVGISSGAAAAAAAFOJLKRPNNAKGLVAVFPPS 300
Db      241  NIIDVEVOISSESETEPAKLLKKEGLLVGISGAAAAAAAKKAKRENNACKLVAVFPPS 300
Cy      301  FGRYLLSSVLFESYVREARSMTPE 324
Db      301  FGRYLLSSVLFESYVREARSMTPE 324
Db      301  FGRYLLSSVLFESYVREARSMTPE 324

RESULT 4
CYKL_BRAJU
ID      CYKL_BRAJU      STANDARD;      PRT;      322 AA.
AC      021733;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DB      Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfhydrylase)
DB      (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL) (OAS-TLa) .
OS      Brassica juncea (leaf mustard) (Indian mustard) .
OS      Brassicota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids II; Brassicales; Brassicaceae; Brassica.
OX      NCBI_TaxID=3707;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Yitlaseo; TISSUE=Root;
RX      MEDLINE=98281577; Pubmed=9620267;
RA      Schaefer H.J., Haag-Korner A., Rausch T.;
RT      "cDNA cloning and expression analysis of genes encoding GSH synthesis
RT      in roots of the heavy-metal accumulator Brassica juncea L.: evidence
RT      for Cd-induction of a putative mitochondrial gamma-Glutamylcysteine
RT      synthetase isoform.";
RL      Plant Mol. Biol. 37:87-97(1998) .
RC      -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +
CC      acetate.
CC      -1- COPACTOR: Pyridoxal phosphate (By similarity) .
CC      -1- PATHWAY: Cysteine biosynthesis.
CC      -1- SUBUNIT: Homodimer (By similarity) .
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity) .
CC      -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-
CC      synthase family.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/comm/ce/
CC      or send an email to license@ebi.ac.uk).
```

DB 3 SRIKADYTELIGNTLVYLNKAGCGVRAAKLEMEPCSSVDKRGFSMISDAKGL 62  
 QY 66 ITPGKSVLEIPSGNTGIGLAFMAAGYKLIITMPASMSLERIILAFGAEVLTDP 125  
 DB 63 IKPESVLEIPSGNTGIGLAFMAAGYKLIITMPASMSVERIILAFGAEVLTDP 122  
 QY 126 KAKGAGVQKABEILAKTPNAYILQGFENPANKHYETTPGELWKSQDGLDAVSGIGT 185  
 DB 123 KAKGAGVQKABEILAKTPNAYILQGFENPANKHYETTPGELWKSQDGLDAVSGIGT 182  
 QY 186 GGTITGAGKYLKEQNPNIKLGVEPVESPVLSGKPGPHKIQIGAGFISVLEVNLD 245  
 DB 183 GGTITGAGKYLKEQNPNIKLGVEPVESPVLSGKPGPHKIQIGAGFISVLEVNLD 242  
 QY 246 VVOISDEALIEFTAKLAKESGLFVGISSGAAAAAFOIAKRENAKGLIYAVPSPGGRY 305  
 DB 243 VVOISDEALIEFTAKLAKESGLFVGISSGAAAAAFOIAKRENAKGLIYAVPSPGGRY 302  
 QY 306 ISSVLPESVYRRASMTPE 324  
 DB 303 LSTVLPESVYRRASMTPE 321

RESULT 5  
 CYK2\_BRAJU STANDARD; PRT; 324 AA.  
 ID CYK2\_BRAJU STANDARD; PRT; 324 AA.  
 AC 023735.  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrolase)  
 DE (O-acetylserine (thiol)-lyase) (CSase) (OAS-TL) (OAS-TL6).  
 OS Brassica juncea (leaf mustard) (Indian mustard).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustosids II; Brassicales; Brassicaceae; Brassica.  
 OC eustosids II; Brassicales; Brassicaceae; Brassica.  
 NCBI\_TaxID=3707;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Vitasso; TISSUE=Root;  
 RA MEDLINE=98281577; PubMed=9620267;  
 RA Schaefer H.J., Haag-Korner A., Rausch T.;  
 RT "Cloning and expression analysis of genes encoding GSH synthesis  
 in roots of the heavy-metal accumulator *Brassica juncea* L.: evidence  
 for co-induction of a putative mitochondrial gamma-glutamylcysteine  
 synthetase isoform.";  
 RT Plant Mol. Biol. 37:87-97(1998).  
 RL Plant Mol. Biol. 37:87-97(1998).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 synthase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: Y10847; CAA71800.1; -  
 DR HSSP: P12674; IONS.  
 DR InterPro: IPR001926; B6 enzyme beta.  
 DR InterPro: IPR001216; Cys\_synthase\_BS.  
 DR InterPro: IPR005859; Cys\_synthase\_BS.  
 DR InterPro: IPR005856; Cys\_synthase\_BS.  
 DR Pfam: PF00291; PALP. 1.  
 DR Trifam: TIGR01139; CysK. 1.

DR TIGRfam: TIGR01136; CysK. 1.  
 DR PROSITE: PS00901; CYS\_SYNTHASE; 1.  
 KW Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.  
 FT Transferring 48 48  
 FT DOMAIN 274 279 POLY-ALA.  
 SQ SEQUENCE 324 AA; 34132 MW; 091B5B2C0BED1EAL CRC64;  
 Query Match 86.8%; Score 1409; DB 1; Length 324;  
 Best Local Similarity 84.1%; Pred. No. 2, 1e-89;  
 Matches 269; Conservative 30; Mismatches 21; Indels 0; Gaps 0;  
 QY 5 RSGIADYTELIGNTLVYLNKAGCGVRAAKLEMEPCSSVDKRGFSMISDAKGL 64  
 DB 4 RSGIADYTELIGNTLVYLNKAGCGVRAAKLEMEPCSSVDKRGFSMISDAKGL 63  
 QY 65 ITPGKSVLEIPSGNTGIGLAFMAAGYKLIITMPASMSLERIILAFGAEVLTDP 124  
 DB 64 ITPGKSVLEIPSGNTGIGLAFMAAGYKLIITMPASMSLERIILAFGAEVLTDP 123  
 QY 125 KAKGAGVQKABEILAKTPNAYILQGFENPANKHYETTPGELWKSQDGLDAVSGIGT 184  
 DB 124 KAKGAGVQKABEILAKTPNAYILQGFENPANKHYETTPGELWKSQDGLDAVSGIGT 183  
 QY 185 GGTITGAGKYLKEQNPNIKLGVEPVESPVLSGKPGPHKIQIGAGFISVLEVNLD 244  
 DB 184 GGTITGAGKYLKEQNPNIKLGVEPVESPVLSGKPGPHKIQIGAGFISVLEVNLD 243  
 QY 245 VVOISDEALIEFTAKLAKESGLFVGISSGAAAAAFOIAKRENAKGLIYAVPSPGGRY 304  
 DB 244 VVOISDEALIEFTAKLAKESGLFVGISSGAAAAAFOIAKRENAKGLIYAVPSPGGRY 303  
 QY 305 ISSVLPESVYRRASMTPE 324  
 DB 304 ISSVLPESVYRRASMTPE 323

RESULT 6  
 CYK2\_ARATH STANDARD; PRT; 322 AA.  
 ID CYK2\_ARATH STANDARD; PRT; 322 AA.  
 AC P47998; O23943; Q42570;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrolase) (O-  
 acetylserine (thiol)-lyase) (CSase A) (CS-A) (OAS-TL A) (Cys-3A)  
 DE (At.OAS-5-8). OR AT4G14880 OR DL3480C.  
 GN OAS1 OR OAS1 OR AT4G14880 OR DL3480C.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Leaf;  
 RA MEDLINE=94364481; PubMed=8082776;  
 RA Hehl R., Bork C., Bogdanova N., Frolov I., Hauschild R.;  
 RT "Isolation and characterization of two cDNAs encoding for compartment  
 specific isoforms of O-acetylserine (thiol) lyase from *Arabidopsis*  
 thaliana.";  
 RT FEBS Lett. 351:257-262(1994).  
 RL FEBS Lett. 351:257-262(1994).  
 CC REVISIONS.  
 RP Hehl R.;  
 RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95246850; PubMed=7729527;  
 RA Barbo C., Vega J.M., Gocer C.;  
 RT "A new member of the cytosolic O-acetylserine (thiol) lyase gene family  
 in *Arabidopsis thaliana*.";  
 RL FEBS Lett. 363:1-5(1995).  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, Columbia;  
 RX MEDLINE=20400348; PubMed=10940562;  
 RT "Genomic and functional characterization of the *cas* gene family  
 encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final  
 step in cysteine biosynthesis in *Arabidopsis thaliana*.";  
 RL Gene 253:237-247(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, Columbia;  
 RX MEDLINE=98121113; PubMed=9461215;  
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,  
 Ridgway P., Dierker W., van Staveren M., Stiekema W., Drost L.,  
 Medley H., Wedler E., Wambutt R., Weitzemagier T., Pohl T., Terry N.,  
 Giehl J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,  
 Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S.,  
 Kottler P., Entian K.-D., Rieger M., Schaefer M., Funk B.,  
 Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,  
 Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,  
 Piravandi E., Obermayer B., Hilbert H., Duesterhoeft A., Moores T.,  
 Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansoorge W.,  
 Cooke R., Berger C., Delzeny M., Voet M., Volckaert G., Mewes H.-W.,  
 Klosterman S., Schueller C., Chaltatzis N.;  
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
*Arabidopsis thaliana*.";  
 RL Nature 391:485-488(1998).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, Columbia;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,  
 Harris B., Ansoorge W., Brandt P., Grievell L.A., Rieger M., Mueller M.,  
 Weitzemagier T., de Simone V., Obermayer B., Macie R., Schmidt T.,  
 Kreis M., Delzeny M., Puigdomenech P., Watson M., Schmidt T.,  
 Reichert B., Portetle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P.,  
 Landham S.-A., McCallagh B., Bilham L., Robben J.,  
 Van der Schueren J., Gylmoprez B., Chuang Y.-J., Vandussche F.,  
 Braeken M., Weijens I., Voet M., Bastiaens I., Bert R., Defoor E.,  
 Weitzemagier T., Botne G., Ramepger U., Hilbert H., Braun M.,  
 Holzer E., Brandt A., Peters S., van Staveren M., Dierker W.,  
 Moollman P., Klein Lankhorst R., Rose M., Hauf J., Kottler P.,  
 Bernier S., Hempel S., Feldpausch M., Lambrecht S., Van den Daele H.,  
 De Keyser A., Buyschaert C., Giehl J., Villarroel R., De Clercq R.,  
 Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,  
 Clark L., Doggett J., Hall S., Kay W., Lennard N., McLeay K., Meyes R.,  
 Reitel A., Rajadream M.A., Lyne M., Benes V., Rechmann S.,  
 Borikova D., Bloecher H., Scharf M., Gimm M., Loehner T.-H.,  
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,  
 Neumann S., Origliou A., Vitale D., Liguori R., Piravandi E.,  
 Massener O., Angley F., Clabaud G., Muendlein A., Felber R.,  
 Schmedl S., Hiller K., Schmidt W., Lecharny A., Aubourg S.,  
 Chetani F., Cooke R., Berger C., Montfort A., Caesacueria E.,  
 Gibbons T., Weber N., Vandembol M., Baryes M., Terol J., Torres A.,  
 Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse C.,  
 Heijlen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,  
 Friedman P., Hase D., Lemcke K., Mewes H.-W., Stocker S.,  
 Zaccaria P., Hasan M., Wilson R.K., de la Bastide M., Habermann K.,  
 Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 Sekhon M., Murray J., Shest P., Cordes M., Abu-Itreiden U.,  
 Stoeckling T., Kallio J., Graves T., Harmon G., Edwards J.,  
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kamp K.,  
 Kramer J., Fulton J., Mardis E., Dante M., Pepin K., Hillier L.W.,  
 Nelson J., Spieth J., Ryan B., Andrews S., Geisel C., Layman D.,  
 Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,  
 Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
 Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,

RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
 Chen R., Marra M.A., Martienssen R., McCombie W.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis  
 thaliana*.";  
 RL Nature 402:769-777(1999).  
 CC -1- CANALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 acetate.  
 CC -1- Cofactor: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X80376; CA556593.2; -;  
 CC EMBL: X84097; CA558893.1; -;  
 CC EMBL: AJ272027; CAB72932.1; -;  
 CC EMBL: Z97337; CAB10267.1; -;  
 CC EMBL: AL161540; CAB78530.1; -;  
 CC PIR: A71412; A71412.  
 CC HSRP: P12674; IONS.  
 CC InterPro: IPR001926; B6, enzyme, beta.  
 CC InterPro: IPR001216; Cys\_synthase\_BS.  
 CC InterPro: IPR005859; Cys\_synth.  
 CC InterPro: IPR005856; Cys\_synthM.  
 CC Pfam: PF00291; PALP; 1.  
 CC TIGRFAMs: TIGR01139; PALP; 1.  
 CC TIGRFAMs: TIGR01136; CysKM; 1.  
 CC PROSITE: PS00801; CYS\_SYNTHASE; 1.  
 CC Transfaser: CysKM; 1.  
 CC Transfaser: CysKM; 1.  
 CC BINDING: 46  
 CC FT DOMAIN: 272 273  
 CC FT CONFLICT: 273 273 A -> E (in REF. 3).  
 CC SEQUENCE: 322 AA; 33805 MW; 5B3E7F3D9DA5908B CRC64;  
 SQ  
 Query Match 85.3%; Score 1384; DB: 1; Length 322;  
 Best Local Similarity 83.4%; Pred. No. 1,1e-87;  
 Matches 266; Conservative 30; Mismatches 23; Indels 0; Gaps 0;  
 QY 6 SGIAADYTELIGTLYLVANKLADGCVARVAKLELAPPCSSVKDRIQYMTADAEKGL 65  
 DB 3 SRIADYVELIGNTPLVYLVANNVABGCVRAKLELAPPCSSVKRIGPSMTSDAEKGL 62  
 QY 66 ITPGKSVLIEPSSGNTGIGLAFMAAARGKYLITPMSLSLERRIILAFGLVLTDP 125  
 DB 63 ITPGKSVLIEPSSGNTGIGLAFMAAARGKYLITPMSLSLERRIILAFGLVLTDP 122  
 QY 126 KKKAKAVOKAABEILAKTNNAYVILQOEPNPKVYETTPGPIWKSQDKIDAFVSGIGT 185  
 DB 123 KKKAKAVOKAABEILAKTNNAYVILQOEPNPKVYETTPGPIWKSQDKIDAFVSGIGT 182  
 QY 186 GGTITGAGKYKEQNPNKILGVEVESPVLSSGFGPGPHKTGIGAGTIPGVLBNVLDE 245  
 DB 183 GGTITGAGKYKEQNPNKILGVEVESPVLSSGFGPGPHKTGIGAGTIPGVLBNVLDE 242  
 QY 246 VVQISDPAIEFPAKLALKEGLFVIGISSGAAALAAAFQIAKRPENAGKIIVAFPSFGERY 305  
 DB 243 VVQISDPAIEFPAKLALKEGLFVIGISSGAAALAAAFQIAKRPENAGKIIVAFPSFGERY 302  
 QY 306 LSSVLFESVRRASQMTPE 324  
 DB 303 LSTVLFDAIRKRAEAMTPE 321  
 RESULT 7

CYSK\_MAIZE  
ID CYSK\_MAIZE STANDARD: PRT; 325 AA.  
AC P80608;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
OS Zea mays (Maize).  
OC Burkholderia viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACAD clade; Panicoidae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. B73; TISSUE=Root;  
RA Brander K.A., Oettrich G.W., Brumold C.;  
RT "Isolation of a cDNA encoding a putative chloroplastic isoform of  
RT cysteine synthase from maize."  
RL (1) Plant Gene Register PGR95-031.  
RN [2]  
RP SEQUENCE OF 11-25.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Danervall C., Huet J.-C.,  
RA Penollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program."  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X85803; CAAS9798.1; -.  
DR PIR: S52738; S52738.  
DR Maize-2DPAGE; P80608; COLBOPTILE.  
DR HSSP: P12674; 1FCJ.  
DR MaizeDB; 123922; -.  
DR InterPro: IPR001926; B6 enzyme beta.  
DR InterPro: IPR001216; Cys\_synthase\_BS.  
DR InterPro: IPR005859; Cys\_synthk.  
DR InterPro: IPR005856; Cys\_synthk.  
DR Pfam: PF00291; PALP; 1.  
DR TIGRFAMs: TIGR01139; cysK; 1.  
DR TIGRFAMs: TIGR01136; cysK; 1.  
DR PROSITE: PS00901; CYS\_SYNTHASE; 1.  
DR Transferrase; Cysteine\_biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49  
FT DOMAIN 275 280  
FT POLY-ALA.  
SQ SEQUENCE 325 AA; 34206 MW; 321369326D7CEBD CRC64;  
Query Match 84.9%; Score 1378; DB 1; Length 325;  
Best Local Similarity 84.3%; Pred. No. 2.8e-87;  
Matches 268; Conservative 22; Mismatches 28; Indels 0; Gaps 0;  
QY 8 IADVTELLIKTLPVYANKLADGCVAVAAKLEMPCCSVKCRIGYATDAEKGILIT 67  
DB 8 IADVTELLIKTLPVYANKLADGCVAVAAKLEMPCCSVKCRIGYATDAEKGILIT 67  
QY 68 PGKSVLEPTSGNTGIGLAFMAAARGKYLITTPASSTERRIILAFGAEVLVTPDAGK 127

DB 68 PGKSVLEPTSGNTGIGLAFMAAARGKYLITTPASSTERRIILAFGAEVLVTPDAGK 127  
QY 128 MGAIVQAEERILATPNAYILQFENPANDPVHYETTPETIKWSDGKIDAFVSGIGTGG 187  
DB 128 MGAIVQAEERILATPNAYILQFENPANDPVHYETTPETIKWSDGKIDAFVSGIGTGG 187  
QY 188 TITAGVLEKONNITLIVGEPVPSVSGRGGPHKIGIGAGTPTGVLVFNLLDEV 247  
DB 188 TITAGVLEKONNITLIVGEPVPSVSGRGGPHKIGIGAGTPTGVLVFNLLDEV 247  
QY 248 QISSDEAIEETAKLALKEGIFVGISSGAAAAAFQIAKRENAKLIIVAFPSGGERYLS 307  
DB 248 QISSDEAIEETAKLALKEGIFVGISSGAAAAAFQIAKRENAKLIIVAFPSGGERYLS 307  
QY 308 SVLPESVRRRAESMTPEP 325  
DB 308 SVLPESVRRRAESMTPEP 325  
QY 308 SVLPESVRRRAESMTPEP 325  
DB 308 SVLPESVRRRAESMTPEP 325  
RESULT 8  
CYK2\_ORYSA  
ID CYK2\_ORYSA STANDARD: PRT; 325 AA.  
AC Q9XEA8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulphydrylase)  
DE (O-acetylserine (Thiol)-lyase) (CSase) (OAS-TL).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Emaritoidae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RX MEDLINE=9196994; PubMed=10095115;  
RA Nakamura T., Yamaguchi Y., Sano H.;  
RT "Four rice genes encoding cysteine synthase: isolation and  
RT differential responses to sulfur, nitrogen and light."  
RL Gene 229:155-161(1999).  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: AF073697; AAD23909.1; -.  
DR HSSP: P12674; 1FCJ.  
DR AND-2DPAGE; Q9XEA8; -.  
DR Gramene; Q9XEA8; -.  
DR InterPro: IPR001926; B6 enzyme beta.  
DR InterPro: IPR001216; Cys\_synthase\_BS.  
DR InterPro: IPR005859; Cys\_synthk.  
DR InterPro: IPR005856; Cys\_synthk.  
DR Pfam: PF00291; PALP; 1.  
DR TIGRFAMs: TIGR01139; cysK; 1.  
DR TIGRFAMs: TIGR01136; cysK; 1.  
DR PROSITE: PS00901; CYS\_SYNTHASE; 1.  
DR Transferrase; Cysteine\_biosynthesis; Pyridoxal phosphate.  
FT BINDING 49 49  
FT POLY-ALA.  
FT PYRIDOXAL PHOSPHATE (By similarity).  
FT

FT DOMAIN 275 280 POLY-ALA.  
 SEQ SEQUENCE 325 AA; 34306 MW; 8F6B0CC0CDD7DCAC4 CRC64;  
 Query Match 84.8%; Score 1377; DB 1; Length 325;  
 Best Local Similarity 80.9%; Pred. No. 3.3e-87;  
 Matches 263; Conservative 28; Mismatches 34; Indels 0; Gaps 0;

QY 1 KAVERSGIAKDYTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVDRIGYSMLADA 60  
 1 MAESGQSIASDVTALIGMTPLVYLNKRVGCEAQAIAKLEIMPCSSVDRIGYSMTIDA 60  
 DB 61 EKKGLITPKSVLIEFTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEV 120  
 61 EKKGLITPKSVLIEFTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEV 120  
 DB 61 EKKGLITPKSVLIEFTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEV 120  
 61 EKKGLITPKSVLIEFTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEV 120  
 QY 121 LTDPKAKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFV 180  
 121 LTDPKAKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFV 180  
 DB 121 LTDPKAKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFV 180  
 121 LTDPKAKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFV 180  
 QY 181 SGIGTCGTTTGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEV 240  
 181 SGIGTCGTTTGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEV 240  
 DB 181 SGIGTCGTTTGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEV 240  
 181 SGIGTCGTTTGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEV 240  
 QY 241 NLDEYVOISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPS 300  
 241 NLDEYVOISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPS 300  
 DB 241 NLDEYVOISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPS 300  
 241 NLDEYVOISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPS 300  
 QY 301 FGERYLSVLFSSVRRRABSMTFEP 325  
 301 FGERYLSVLFSSVRRRABSMTFEP 325  
 DB 301 FGERYLSVLFSSVRRRABSMTFEP 325  
 301 FGERYLSVLFSSVRRRABSMTFEP 325

RESULT 9  
 CYSK WHEAT STANDARD; PRT; 325 AA.  
 AC P38076;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
 DB (O-acetylserine (Thiol)-lyase) (Cbase A) (OAS-TL A).  
 GN CYS1.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 OC Triticaceae; Triticum.  
 NC NCB1\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=94100990; PubMed=8275097;  
 RA Yousefian S., Nakamura M., Sano H.;  
 RT "Tobacco plants transformed with the O-acetylserine (thiol) lyase  
 RT gene of wheat are resistant to toxic levels of hydrogen sulphide  
 RT gas.";  
 RT Plant J. 4:759-769 (1993).  
 RL -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 RL acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -1- SUBUNIT: Homodimer (by similarity).

DR EMBL; D13153; BA02438.1; -;  
 DR PIR; J50762; J50762.  
 DR HSSP; P12674; 10AS.  
 DR InterPro; IPR001926; B6 enzyme beta.  
 DR InterPro; IPR001216; Cys synthase BS.  
 DR InterPro; IPR005859; Cys synthase.  
 DR InterPro; IPR005856; Cys synthase.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; CysK; 1.  
 DR TIGRFAMs; TIGR01136; CysK; 1.  
 DR PROSITE; PS00901; CYS SYNTHASE; 1.  
 KW Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.  
 FT BINDING 50 50 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT DOMAIN 276 281 POLY-ALA.  
 FT SEQUENCE 325 AA; 34114 MW; 1CEFBFE29F341DCA CRC64;  
 Query Match 84.0%; Score 1364; DB 1; Length 325;  
 Best Local Similarity 82.3%; Pred. No. 2.5e-86;  
 Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKDYTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVDRIGYSMLADAERKGLIT 67  
 8 IAKDYTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVDRIGYSMLADAERKGLIT 67  
 DB 9 IAKDYTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVDRIGYSMTIDAEKGIIV 68  
 9 IAKDYTELIGKTPLYLNLKADGCVARVAKLELMEPCSSVDRIGYSMTIDAEKGIIV 68  
 QY 68 FGSYVLIPTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEVLTDPKAG 127  
 68 FGSYVLIPTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEVLTDPKAG 127  
 DB 69 FGSYVLIPTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEVLTDPKAG 128  
 69 FGSYVLIPTSGNTGIGLAFMAAGCYKLIITMPASMSLERIILARAEVLTDPKAG 128  
 QY 128 MKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFVSGITGCG 187  
 128 MKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFVSGITGCG 187  
 DB 129 MKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFVSGITGCG 188  
 129 MKGAVOKABEILAKTPNAYILQOFENPANPKVYETTGPEIMKSGDGKIDAFVSGITGCG 188  
 QY 188 TITGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEVNLDEYV 247  
 188 TITGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEVNLDEYV 247  
 DB 189 TITGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEVNLDEYV 248  
 189 TITGAKTILKEONPNIKLIGVPVPSVULSGKGPPEHKIIGAGFTTGVLEVNLDEYV 248  
 QY 248 QISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPSGERYLS 307  
 248 QISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPSGERYLS 307  
 DB 249 QISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPSGERYLS 308  
 249 QISSDEATEAKTLAKBGLFVGISGGAIAAALAFOLAKPENAKGLIYAVFPSGERYLS 308  
 QY 308 SVLFSSVRRRABSMTFEP 324  
 308 SVLFSSVRRRABSMTFEP 324  
 DB 309 SVLFSSVRRRABSMTFEP 325  
 309 SVLFSSVRRRABSMTFEP 325

RESULT 10  
 CYKL ORYZA STANDARD; PRT; 321 AA.  
 ID CYKL\_ORYZA  
 AC Q9XKA6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase (EC 2.5.1.47) (O-acetylserine sulfinhydrilase)  
 DB (O-acetylserine (Thiol)-lyase) (Cbase A) (OAS-TL).  
 GN RGS1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhacoidae; Oryzae; Oryza.  
 NC NCB1\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RX MEDLINE=99196994; PubMed=10095115;  
 RA Nakamura T., Yamaguchi Y., Sano H.;  
 RT "Four rice genes encoding cysteine synthase: isolation and  
 RT differential responses to sulfur, nitrogen and light.";  
 RT Gene 229:155-161 (1999).  
 RL -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 RL acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate (by similarity).  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer (by similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF073695; AAD23907.1; -.  
CC HSSP; P12674; IFCU.  
CC Gramene; Q9XEA6; -.  
CC InterPro; IPR001926; B6\_enzyme\_beta.  
CC InterPro; IPR001216; Cys\_synthase\_BS.  
CC InterPro; IPR005859; Cys\_synthase.  
CC InterPro; IPR005856; Cys\_synthase.  
CC Pfam; PF00291; PALP; 1.  
CC TIGRfam; TIGR01139; cysK; 1.  
CC TIGRfam; TIGR01136; cysK; 1.  
CC PROSITE; PS00901; CYS\_SYNTHASE; 1.  
CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate.  
CC BINDING; 46  
CC FT DOMAIN 272 277 POLY-ALA.  
CC FT SEQUENCE 321 AA; 3386 MW; 9DCC2A2C9A6D63A6 CRC64;  
SQ  
Query Match 83.5%; Score 1356; DB 1; Length 321;  
Best Local Similarity 81.7%; Pred. No. 8,8e-86;  
Matches 259; Conservative 32; Mismatches 26; Indels 0; Gaps 0;  
QY 8 IADQVLELIGKPLVYLNKADGCVARVAKLELMEPCSSVDRIGYSMTADAEKGLT 67  
DB 5 IADQVLELIGKPLVYLNKADGCVARVAKLELMEPCSSVDRIGYSMTADAEKGLT 64  
QY 68 PGKSVLEPTSGNTGIGLAFMAAARGYKLTITPASPMSLERRIILAFGAELVLDPAAG 127  
DB 65 PGKSVLEPTSGNTGIGLAFMAAARGYKLTITPASPMSLERRIILAFGAELVLDPAAG 124  
QY 128 MKANQKAEELIKTPAVYILQFENPANKVHYETTPRIKSGDKIDAFVSGIGTGS 187  
DB 125 MKANQKAEELIKTPAVYILQFENPANKVHYETTPRIKSGDKIDAFVSGIGTGS 184  
QY 188 TITGAGKYLEKQNPNIKILIGVEPVESPVLSGGRPGPKIGIGAGF1PGYLEVNLDEVV 247  
DB 185 TITGAGKYLEKQNPNIKILIGVEPVESPVLSGGRPGPKIGIGAGF1PGYLEVNLDEVV 244  
QY 248 QIISDAELETAKLAKKEGFPVIGISSGAAAAAFAQAKRENNAGKILVAFPSRGKYS 307  
DB 245 QIISDAELETAKLAKKEGFPVIGISSGAAAAAFAQAKRENNAGKILVAFPSRGKYS 304  
QY 308 SVLFESVYRREASNTFE 324  
DB 305 SVLFESVYRREASNTFE 321

RP SEQUENCE FROM N.A.  
RC STRAIN-CEV. Bercolina;  
RA Heese H., Hoelgen R.;  
RT "Isolation of cDNAs encoding cytosolic and plastidic cysteine synthase  
CC isoforms from Solanum tuberosum."  
CC (in) Plant Gene Register PCR98-057.  
CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
CC acetate.  
CC -1- Cofactor: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Cysteine biosynthesis.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
CC synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF044173; AAC25636.1; -.  
CC PIR; T07002; T07002.  
CC HSSP; P12674; 10AS.  
CC InterPro; IPR001926; B6\_enzyme\_beta.  
CC InterPro; IPR001216; Cys\_synthase\_BS.  
CC InterPro; IPR005859; Cys\_synthase.  
CC InterPro; IPR005856; Cys\_synthase.  
CC Pfam; PF00291; PALP; 1.  
CC TIGRfam; TIGR01139; cysK; 1.  
CC TIGRfam; TIGR01136; cysK; 1.  
CC PROSITE; PS00901; CYS\_SYNTHASE; 1.  
CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
CC Transic peptide.  
CC TRANSIT 52  
CC FT CHAIN 53 386 CYSTEINE SYNTHASE.  
CC FT BINDING 110 110 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC SQ SEQUENCE 386 AA; 41044 MW; 8EB3FEB2084D9594 CRC64;  
Query Match 81.0%; Score 1314; DB 1; Length 386;  
Best Local Similarity 78.1%; Pred. No. 8e-83;  
Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;  
QY 3 VERSGIADVLELIGKPLVYLNKADGCVARVAKLELMEPCSSVDRIGYSMTADAE 62  
DB 64 IEGNLIAEDVDTQIGNTFVYLNKADGCVARVAKLELMEPCSSVDRIGYSMTADAE 123  
QY 63 KELTPGKSVLEPTSGNTGIGLAFMAAARGYKLTITPASPMSLERRIILAFGAELVLT 122  
DB 124 KELTPGKSVLEPTSGNTGIGLAFMAAARGYKLTITPASPMSLERRIILAFGAELVLT 183  
QY 123 DPAKMKGAQVAKAEILAKTPAVYILQFENPANKVHYETTPRIKSGDKIDAFVSG 182  
DB 184 DPAKMKGAQVAKAEILAKTPAVYILQFENPANKVHYETTPRIKSGDKIDAFVSG 243  
QY 183 IGTGAGKYLEKQNPNIKILIGVEPVESPVLSGGRPGPKIGIGAGF1PGYLEVNL 242  
DB 244 IGTGAGKYLEKQNPNIKILIGVEPVESPVLSGGRPGPKIGIGAGF1PGYLEVNL 303  
QY 243 LDEVAQISSDAELETAKLAKKEGFPVIGISSGAAAAAFAQAKRENNAGKILVAFPSRG 302  
DB 304 LDEVAQISSDAELETAKLAKKEGFPVIGISSGAAAAAFAQAKRENNAGKILVAFPSRG 363  
QY 303 ERYLSSVLFESVYRREASNTFE 321  
DB 364 ERYLSSVLFESVYRREASNTFE 382

RESULT 11  
CYSL\_SOLTU STANDARD; PRT; 386 AA.  
ID 081155;  
AC 16-OCT-2001 (Rel. 40, last sequence update)  
DT 16-OCT-2001 (Rel. 40, last sequence update)  
DE 10-OCT-2003 (Rel. 42, last annotation update)  
DE Cysteine synthase, chloroplast precursor (BC 2.5.1.47) (O-acetylserine  
DE sulfhydrylase) (O-acetylsulfhydrylase) (CSase B) (CS-B) (OAS-B)  
DE B).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]

RESULT 12  
CYSL\_SPIOL

ID CYSL SP10L STANDARD; PRT; 383 AA.  
 AC P33260; 033137;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, chloroplast precursor (EC 2.5.1.47) (O-acetylserine  
 sulfhydrilase) (O-acetylserine (thiol)-lyase) (CSase B) (CS-B) (OAS-TL  
 B).  
 GN CysK.  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 NC NCB1\_TaxID=3562;  
 RX MEDLINE=94403590; PubMed=8405359;  
 RA Saito K., Tateyuchi K., Murakoshi I., Hirano H.;  
 RT "cDNA cloning and expression of cysteine synthase B localized in  
 chloroplasts of Spinacia oleracea.";  
 RL FEBS Lett. 324:247-252(1993).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=93143317; PubMed=8424655;  
 RA Rolland N., Droux M., Lebun M., Douce R.;  
 RT "O-acetylserine(thiol)lyase from spinach (Spinacia oleracea L.) leaf:  
 cDNA cloning, characterization, and overexpression in Escherichia  
 coli of the chloroplast isoform.";  
 RL Arch. Biochem. Biophys. 300:213-222(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Marathon; TISSUE=leaf;  
 RX MEDLINE=94105326; PubMed=8278530;  
 RA Hall R., Schuster G., Griseem M.;  
 RT "An O-acetylserine (thiol) lyase cDNA from spinach.";  
 RL Plant Physiol. 102:1057-1058(1993).  
 RN [4]  
 RP DISCUSSION OF SEQUENCE.  
 RX MEDLINE=93356752; PubMed=7916619;  
 RA Rolland N., Job D., Douce R.;  
 RT "Common sequence motifs coding for higher-plant and prokaryotic O-  
 acetylserine (thiol)-lyases: bacterial origin of a chloroplast  
 transist peptide?";  
 RL Biochem. J. 293:829-833(1993).  
 RN [5]  
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=96184908; PubMed=8617276;  
 RA Rolland N., Ruffet M.-L., Job D., Douce R., Droux M.;  
 RT "Spinach chloroplast O-acetylserine (thiol)-lyase exhibits two  
 catalytically non-equivalent pyridoxal-5'-phosphate-containing active  
 sites.";  
 RL Eur. J. Biochem. 236:272-282(1996).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chloroplast.  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: D14722; BAA03542.1; -;  
 CC EMBL: X66860; CAA47329.1; -;  
 CC EMBL: U05184; AAA16973.1; -;

DR PIR: S29733; S29733.  
 DR PIR: T09000; T09000.  
 DR HSSP: P12674; 1FCJ.  
 DR InterPro: IPR001926; B6 enzyme beta.  
 DR InterPro: IPR001216; Cys synthase\_B6.  
 DR InterPro: IPR005859; Cys synthase.  
 DR InterPro: IPR005859; Cys synthase.  
 DR Pfam: PF00291; PALP; 1.  
 DR TIGRfams: TIGR01139; CysK; 1.  
 DR TIGRfams: TIGR01136; CysK; 1.  
 DR PROSITE: PS00901; CYS SYNTHASE; 1.  
 KW Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
 KM Transist peptide.  
 FT TRANSIT 1 52  
 FT CHAIN 53 383  
 FT BINDING 108 108  
 FT DOMAIN 334 339  
 FT CONFLICT 12 12  
 FT CONFLICT 166 166  
 FT CONFLICT 252 252  
 FT CONFLICT 274 280  
 FT CONFLICT 334 335  
 FT CONFLICT 334 335  
 FT CONFLICT 379 380  
 FT CONFLICT 383 383  
 SQ SEQUENCE 383 AA; 40637 MM; CE3725796D3A3F93 CRC64;  
 Query Match 78.3%; Score 1271; DB 1; Length 383;  
 Best Local Similarity 73.4%; Pred. No. 6.9e-80;  
 Matches 234; Conservative 52; Mismatches 33; Indels 0; Gaps 0;  
 QY 3 VERSGIADVTELGKTPVLYLNKADGCVARVAKELMPPCSSVDORIGYSMTADAE 62  
 DB 62 EKGNTADVDSQIGKTPMYLIANNVSGVANIADKESNEPCQSVDRIGYSMTADAE 121  
 QY 63 KGLITPKSVLIEFTSGNTGIGLAFMAAANGYKILTMPMSLSIERKILLAFSAELVLT 122  
 DB 122 KGVITPKSVLIEFTSGNTGIGLAFMAAANGYKILTMPMSLSIERKILLAFSAELVLT 181  
 QY 123 DPAGKMGAVOKAEIILAKTPNAYIIQFENPANKVHYETTGPEIMKSGDKIDAFVSG 182  
 DB 182 DPAGKMGAVOKAEIILAKTPNAYIIQFENPANKVHYETTGPEIMKSGDKIDAFVSG 241  
 QY 183 IGTGGITGAGKYLKQNPNIKILGVEPVSSPVLSGKPGPHKIQGIAGGFIPEVLEVNL 242  
 DB 242 IGTGGITGAGKYLKQNPNIKILGVEPVSSPVLSGKPGPHKIQGIAGGFIPEVLEVNL 301  
 QY 243 LDEVVOISSDPAITATLALKEGLFVGISGGAAMAAAFQAKPENNAKGLIVAPPSFG 302  
 DB 302 MDEVTESSSEBAVMAQALAKBELLVISGAAAPAAAVRIGKPPENNAKGLIVAPPSFG 361  
 QY 303 ERYLSVLFESVREAEASM 321  
 DB 362 ERYLSVLFESVREAEASM 380  
 RESULT 13  
 CYSM ARATH STANDARD; PRT; 424 AA.  
 AC Q43725; QSM128; QSM440;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, mitochondrial precursor (EC 2.5.1.47) (O-  
 acetylserine sulfhydrilase) (O-acetylserine (thiol)-lyase) (CSase C)  
 (CS-C) (OAS-TL C) (ATCS-C).  
 GN OAS1 OR ACS1 OR AT3G59760 OR F24G16.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Euphorbiales; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCB1\_TaxID=3702;  
 RX MEDLINE=93356752; PubMed=7916619;  
 RA Rolland N., Job D., Douce R.;  
 RT "Common sequence motifs coding for higher-plant and prokaryotic O-  
 acetylserine (thiol)-lyases: bacterial origin of a chloroplast  
 transist peptide?";  
 RL Biochem. J. 293:829-833(1993).







RA Hell R., Bork C., Bogdanova N., Frolov I., Hauschild R.;  
 RT "Isolation and characterization of two cDNAs encoding for compartment  
 RT specific isoforms of O-acetylserine (thiol) lyase from Arabidopsis  
 RT thaliana.";  
 RT FEBS Lett. 351:257-262(1994).  
 RM [2]  
 RM REVIEWS.  
 RM Hell R.;  
 RM Submitted (XXX-2000) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=95334513; PubMed=7610184;  
 RA Hesse H., Altman T.;  
 RT "Molecular cloning of a cysteine synthase cDNA from Arabidopsis  
 RT thaliana.";  
 RL Plant Physiol. 108:851-852(1995).  
 RM [4]  
 RM SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20400348; PubMed=10940562;  
 RA Jost R., Berkowitz O., Wirtz M., Hopkins L., Hawkesford M.J., Hell R.;  
 RT "Genomic and functional characterization of the oae gene family  
 RT encoding O-acetylserine (thiol) lyases, enzymes catalyzing the final  
 RT step in cysteine biosynthesis in Arabidopsis thaliana.";  
 RL Gene 253:237-247(2000).  
 RM [5]  
 RM SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.-J., Renning C.M., Koo H.L.,  
 RA Moffat K.S., Cronin L.A., Shen M., Pat G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.B., Adams M.D., Carrera A.V., Creasy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast.  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 CC EMBL: X80377; CAA56594.2; -  
 CC EMBL: X81698; CAA57344.1; -  
 CC EMBL: AJ271728; CAB71297.1; -  
 CC EMBL: AC002333; AAB64031.1; -  
 CC PIR: A84870; A84870.  
 CC PIR: S48695; S48695.  
 CC HSSP: P12674; 10AS.  
 CC SWISS-2DPAGE: P47999; ARATH.  
 CC InterPro: IPR001926; B6 enzyme beta.  
 CC InterPro: IPR001216; Cys synthase beta.  
 CC InterPro: IPR005859; Cys synthase.  
 CC InterPro: IPR005856; Cys synthase.  
 CC Pfam: PF00291; PALP; 1.  
 CC TIGRfam: TIGR01139; CysK; 1.  
 CC TIGRfam: TIGR01136; CysK; 1.

DR PROSITE, PS00901; CYS SYNTHASE; 1.  
 KM Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
 KM Transf. peptide.  
 FT TRANSIT 1 60 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 61 392 CYSTEINE SYNTHASE.  
 FT BINDING 116 116 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT DOMAIN 342 347 POLY-ALA.  
 FT CONFLICT 232 232 G -> N (IN REF. 3).  
 FT CONFLICT 288 288 P -> T (IN REF. 3).  
 SQ SEQUENCE 392 AA; 41656 MM; B1220F01E831A12 CMC64;  
 Query Match 77.7%; Score 1261; DB 1; Length 392;  
 Best local similarity 74.9%; Pred. No. 3,4e-79;  
 Matches 239; Conservative 41; Mismatches 39; Indels 0; Gaps 0;  
 QY 3 VERSGIADVTETLIGKPLVYINLADCVARVAKLEMEPCSSVDRIQYIMADAE 62  
 DB 70 VEGNIADNMAQLIGKTPMYINNVGCVASVAKLEIMEPCSSVDRIQYIMTDAEE 129  
 QY 63 KGLITPGKSVLIPRTSGNTGIGLAPMAARGKILITPMSLERRITLLARAEVLT 122  
 DB 130 KGLITPGKSVLIPRTSGNTGIGLAPMAARGKILITPMSLERRITLLARAEVLT 189  
 QY 123 DPAKMGKAVOKASEILAKTPNAYILQOFENPANKVHYETTGPBEIMKSGDKIDAFVSG 182  
 DB 190 EPAKMGKAVOKASEILAKTPNAYILQOFENPANKVHYETTGPBEIMKSGDKIDAFVSG 249  
 QY 183 IGTGTTGAGKVIKAEQNPINKLIGVPVSPVLSGKPGPHKIQGIGAGTIPGLVNL 242  
 DB 250 IGTGTTGAGKVIKAEQNPINKLIGVPVSPVLSGKPGPHKIQGIGAGTIPGLVNL 309  
 QY 243 IDEVVOISDPAIFAKILAKKEGLFVGISGAAAFAPOLAKPENAGKLIIVPSPFG 302  
 DB 310 VDEVVOISDPAIFAKILAKKEGLFVGISGAAAFAPOLAKPENAGKLIIVPSPFG 369  
 QY 303 ERYLSSVLPESVREAEEM 321  
 DB 370 ERYLSSVLPESVREAEEM 388  
 RESULT 15  
 ID CYSL CAPAN STANDARD; PRT; 374 AA.  
 AC P31300;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine synthase, chloroplast precursor (EC 2.5.1.47) (O-acetylserine  
 DE sulfhydrylase) (O-acetylserine (thiol)-lyase) (Case B) (CS-B) (ONS-TL  
 DE B).  
 OS Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Capsicum.  
 OX NCBI\_TaxID=4072;  
 RM [1]  
 RM SEQUENCE FROM N.A.  
 RC STRAIN=cv. Lamuyo; TISSUE=fruit;  
 RX MEDLINE=92389158; PubMed=1381358;  
 RA Roemer S., D'Harlingue A., Camara B., Schantz R., Kuntz M.;  
 RT "Cysteine synthase from Capsicum annuum chromoplasts.  
 RT Characterization and cDNA cloning of an up-regulated enzyme during  
 RT fruit development.";  
 RL J. Biol. Chem. 267:17966-17970(1992).  
 CC -1- CATALYTIC ACTIVITY: O(3)-acetyl-L-serine + H(2)S = L-cysteine +  
 CC acetate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- PATHWAY: Cysteine biosynthesis.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast.  
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-  
 CC synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
CC EMBL; X64874; CAA46086.1; -  
CC PIR; A43407; A43407.  
CC HSP; P12674; 10AS.  
CC InterPro; IPR001926; B6 enzyme beta.  
CC InterPro; IPR001216; Cys\_synthase\_BS.  
CC InterPro; IPR005859; Cys\_synthase\_BS.  
CC InterPro; IPR005856; Cys\_synthase\_BS.  
CC Pfam; PF00291; PALP; 1.  
CC TIGRfams; TIGR01139; cysK; 1.  
CC TIGRfams; TIGR01136; cysK; 1.  
CC PROSITE; PS00901; CYS SYNTHASE; 1.  
CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
CC Transferrase; Cysteine biosynthesis; Pyridoxal phosphate; Chloroplast;  
CC TRANSIT peptide.  
CC TRANSIT 1 50 CHLOROPLAST (BY SIMILARITY).  
CC CHAIN 51 374 CYSTEINE SYNTHASE.  
CC BINDING 108 108 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC SEQUENCE 374 AA; 39978 MW; 36A6D0CCEA23F1 CRC64;

Query Match 77.0%; Score 1249; DB 1; Length 374;

Best Local Similarity 75.2%; Pred. No. 2.1e-78;  
Matches 240; Conservative 38; Mismatches 31; Indels 10; Gaps 1;

QY 3 VERSGIADVTLLIGKTPVLYINKLADGCVARVAKLEMEPCSSVDRIGYSMTADAE 62  
DB 62 IEGNIAEDVTQLIGNTPMVYINTIVKGCVANIAKLEIMEPCSSVDRIGYSMTADAE 121  
QY 63 KGLITPKSVLIERTSGNTGIGAPMAARGKLTITPASMTERIIILAPGARLVT 122  
DB 122 KGLISPGTVAVEPTSGNTGIGAPMAARGKLTITPASMTERIIILAPGARLVT 181  
QY 123 DPAKMGKAVQKAEELIAKTPNAYILQGFENPANKVYETTPGPIWKSDDKIDAFVSG 182  
DB 182 DPAKMGKAVSKAEELIANTPDAYILQGFENPANKVYETTPGPIWKSDDKIDAFVSG 241  
QY 183 IGTGTTGAGKYLKEQPNKILIGVEVESPVLSSGKPGPHKIQIGAGTIPGVLEVN 242  
DB 242 IGTGTTGAGKYLKEQPNKILIGVEVESPVLSSGKPGPHKIQIGAGTIPGVLEVN 291  
QY 243 IDEVVOISSDEAITAKLLAKKEGFWGSSGAAAAAFOIAPKPPENAGKLIIVFPSPG 302  
DB 292 MDEVTEISSDEAITAKLLAKKEGFWGSSGAAAAAFOIAPKPPENAGKLIIVFPSPG 351  
QY 303 ERYLSSVLFESVRRPESM 321  
DB 352 ERYLSSVLFESVRRPESM 370

Search completed: May 18, 2004, 13:07:34  
Job time : 18 secs

Mon May 24 08:18:54 2004

us-09-931-457a-31.rspc

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using SW model

Run on: May 18, 2004, 13:03:34 ; Search time 45 Seconds  
(without alignment)  
2278.743 Million cell updates/sec

Title: US-09-931-457A-31  
Perfect score: 1623  
Sequence: 1 MAVESGIAKDYTELIGKTP.....ISSVLFESVREARESGMTPEP 325

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_prokaryote:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1622	99.9	325	10	Q8W1A0 glycine max
2	1445	89.0	325	10	Q9F827 solanum tub
3	1373	84.6	325	10	Q9MA22 allium tub
4	1366.5	84.2	321	10	Q94AS7 arabidopsis
5	1316	81.1	366	10	Q9F826 solanum tub
6	1296	79.5	305	10	Q9L1A0 arabidopsis
7	1290	78.6	266	10	Q65747 cicer ariet
8	1275.5	78.6	409	10	Q845B4 oryza sativ
9	1271	78.3	430	10	Q814A2 arabidopsis
10	1270	78.3	352	10	Q81523 chlamydomon
11	1246	76.8	376	10	Q9F825 solanum tub
12	1229	75.7	324	10	Q43726 arabidopsis
13	1227	75.6	282	10	Q8W1B3 oryza sativ
14	1220	75.2	324	10	Q9F8F5 nicotiana t
15	1192	73.4	323	10	Q9S5S7 arabidopsis
16	1192	73.4	323	10	Q8LEC3 arabidopsis

17	1184	73.0	324	10	Q8LA06 arabidopsis
18	1179	72.6	324	10	Q9S6Z7 arabidopsis
19	1179	72.6	324	10	Q9CAY8 arabidopsis
20	1161	71.5	345	10	Q23734 brassica ju
21	1086	66.9	389	10	Q9SSV9 arabidopsis
22	1069	65.9	377	10	Q7XBS5 porphyra pu
23	1063	65.5	377	10	Q7XBS6 oryza sativ
24	1061	65.4	377	10	Q9FSL8 oryza sativ
25	1049	64.6	352	10	Q7Y256 betula vert
26	1047	64.5	368	10	Q43153 spinacia ol
27	1017	62.7	368	10	Q9S757 arabidopsis
28	1015	62.5	347	10	Q9F829 solanum tub
29	1004	61.9	319	16	Q8YU40 anabaena sp
30	1001	61.7	321	16	Q8DL17 synechococc
31	989	60.9	320	16	Q8YU13 anabaena sp
32	968.5	59.7	315	16	Q8A375 bacteroides
33	968	59.6	344	5	Q93244 caenorhabdi
34	947	58.3	322	16	Q8U113 agrobacteri
35	945	58.2	322	16	Q92SN1 rhizobium m
36	934	57.5	326	16	Q98DP1 rhizobium l
37	927	57.1	357	10	Q9XEA7 oryza sativ
38	925	57.0	337	5	Q45679 caenorhabdi
39	923	56.9	322	16	Q7VE70 prochloroco
40	919	56.6	322	16	Q7V4H3 prochloroco
41	916	56.4	336	16	Q8YU12 bruceella m
42	912	56.2	310	16	Q9JUL6 neisseria m
43	909.5	56.0	310	16	Q8KE19 chlorobium
44	902	55.6	339	10	Q9XEA9 oryza sativ
45	898.5	55.4	310	2	Q9F8A6 selenomonas

## ALIGNMENTS

### RESULT 1

Q8W1A0 PRELIMINARY; PRT; 325 AA.

AC Q8W1A0; 01-MAR-2002 (TREMBL) 20, Created)  
DT 01-MAR-2002 (TREMBL) 20, Last sequence update)  
DT 01-OCT-2003 (TREMBL) 25, Last annotation update)  
DB Cysteine synthase.  
OS Glycine max (Soybean).  
OC Burkholderia; Vibrionaceae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC eudicotyledons; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT demosthenis C. Krishnan H.B.;  
RT "Nucleotide sequence of a soybean cDNA encoding cysteine synthase";  
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF452451; AAL6291.1; -  
DR GO: GO:0004124; F:cysteine synthase activity; IEA.  
DR GO: GO:0016829; F:lyase activity; IEA.  
DR GO: GO:0006520; P:amino acid metabolism; IEA.  
DR GO: GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro: IPR001926; B6 enzyme beta.  
DR InterPro: IPR001216; Cys\_synthase\_BS.  
DR InterPro: IPR005859; Cys\_synthase.  
DR InterPro: IPR005856; Cys\_synthase.  
DR Pfam: PF00291; PALP; 1.  
DR TIGRPFAM: TIGR01139; cysK; 1.  
DR TIGRPFAM: TIGR01136; cysK; 1.  
DR PROSITE: PS00901; CYS\_SYNTHASE; 1.  
SQ SEQUENCE 325 AA; 34269 MW; 6834B6F65A03PCC CRC64;

Query Match 99.9%; Score 1622; DB 10; Length 325;  
Best Local Similarity 99.7%; Pred. No. 7.7e-106;  
Matches 324; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
1 MAVESGIAKDYTELIGKTPVYLNKADGCVAAKLEMLPSSVDRIGYSITADA 60

DB 1 MAVERSGIADVTTELIGKTEPLVYLNKLDGCVARVAKLEMEPCSSVKDRIGYSMTADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALV 120  
 DB 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALV 120  
 QY 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180  
 DB 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180  
 QY 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEY 240  
 DB 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEY 240  
 QY 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPS 300  
 DB 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPS 300  
 QY 301 FGERIYSSVLPESVRRASNTPEP 325  
 DB 301 FGERIYSSVLPESVRRASNTPEP 325

## RESULT 2

Q9FS27 PRELIMINARY; PRT; 325 AA.

ID Q9FS27  
 AC Q9FS27  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Cytosolic cysteine synthase.  
 GN PCS-1.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxId=4113;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA STRAIN=Dansyaku.  
 RA Maruyama A., Ishizawa K.;  
 RT "Cytosolic Cysteine Synthase from Potato Sprout."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB029511; BAB20861.1; -.  
 DR HSSP; P12674; IOAS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6 enzyme beta.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 SQ SEQUENCE 325 AA; 34308 MW; 980D526562F6C785 CRC64;

Query Match 89.0%; Score 1445; DB 10; Length 325;  
 Best Local Similarity 85.8%; Pred. No. 2e-93;  
 Matches 279; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

QY 1 MAVERSGIADVTTELIGKTEPLVYLNKLDGCVARVAKLEMEPCSSVKDRIGYSMTADA 60  
 DB 1 MAVERSGIADVTTELIGKTEPLVYLNKLDGCVARVAKLEMEPCSSVKDRIGYSMTADA 60  
 QY 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALV 120  
 DB 61 EEKGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALV 120  
 QY 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180  
 DB 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180  
 QY 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180  
 DB 121 LTPDPAKMGKAVQAEELIATKTNAYILQGFENPANKVHYETTGPEIWKSGDKIDAFV 180

QY 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEY 240  
 DB 181 SGIGTGTTGAGKYLKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEY 240  
 QY 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPS 300  
 DB 241 NLIDEVVOISSDEALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPS 300  
 QY 301 FGERIYSSVLPESVRRASNTPEP 325  
 DB 301 FGERIYSSVLPESVRRASNTPEP 325

## RESULT 3

Q9MA22 PRELIMINARY; PRT; 325 AA.

ID Q9MA22  
 AC Q9MA22  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Cysteine synthase.  
 GN BSAL.  
 OS Allium tuberosum (Garlic chives).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;  
 OC Allium.  
 OC NCBI\_TaxId=4683;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA MEDLINE=20534797; PubMed=11080593;  
 RA Urano Y., Manabe T., Noji M., Saito K.;  
 RT "Molecular cloning and functional characterization of cDNAs encoding  
 RT cysteine synthase and serine acetyltransferase that may be responsible  
 RT for high cellular cysteine content in Allium tuberosum.";  
 RL Gene 257:269-277(2000)  
 DR EMBL; AB040503; BAA93051.1; -.  
 DR HSSP; P12674; IOAS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6 enzyme beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthKM.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRfams; TIGR01139; cysK; 1.  
 DR TIGRfams; TIGR01136; cysKM; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASB; 1.  
 SQ SEQUENCE 325 AA; 34337 MW; C86BD91205DF358F CRC64;

Query Match 84.6%; Score 1373; DB 10; Length 325;  
 Best Local Similarity 82.4%; Pred. No. 2.3e-88;  
 Matches 262; Conservative 32; Mismatches 24; Indels 0; Gaps 0;

QY 8 IADVTTELIGKTEPLVYLNKLDGCVARVAKLEMEPCSSVKDRIGYSMTADAENGKIT 67  
 DB 8 IADVTTELIGKTEPLVYLNKLDGCVARVAKLEMEPCSSVKDRIGYSMTAERKGLIT 67  
 QY 68 PKGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALVITDPAG 127  
 DB 68 PKGKSVLIEPTSGNTGIGLAFMAARGYKLIITMPASMSIERRIILAFAGALVITDPAG 127  
 QY 128 MKGACERAKABIAKTNSVYLOGFENPANKVHYETTGPEIWKSGDKIDAFVSGITG 187  
 DB 128 MKGACERAKABIAKTNSVYLOGFENPANKVHYETTGPEIWKSGDKIDAFVSGITG 187  
 QY 188 TITGAGKYLNKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEYNLIDEV 247  
 DB 188 TITGAGKYLNKEQNPNIKIGVEPVSPTVSGKPGPHKIQGGAGTIPGVLEYNLIDEV 247  
 QY 248 QISSDALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPSGERRYS 307  
 DB 248 QISSDALETAKLALKEGLFVGISGAAAAAFOIAKPEENAGKLIIVAFPSGERRYS 307

Mon May 24 08:18:54 2004

us-09-931-457a-31.rsp

Page 3

DB 248 QVSSSEAIEMAKQALQGLVIGISSGAAALAIKFAVRPENEGHLYVFPSPGERTLS 307  
QY 308 SVLPESVREARASMTFEP 325  
DB 308 TVLPFSIKKAEASMTFEP 325

RESULT 4

Q94AS7 PRELIMINARY; PRT; 321 AA.

AC Q94AS7  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative cytosolic O-acetylserine(thiol)lyase.  
GN AF4G14880.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC euroids II; Brassicales; Brassicaceae; Arabidopsie.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,  
Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,  
Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
Hayashizaki Y., Ishida J., Jones T., Kamiya A., Katlin-Neumann G.,  
Kawai J., Kim C., Koese B., Lam B., Lin J., Meyers M.C., Miranda M.,  
Rakusana M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,  
Shum P., Southwick A., Tracy S.E., Shinzaki K., Davis R.M.,  
Ritter J.R., Theologis A.,  
RT "Full length cDNA of gene AT4G14880 (GI\_7268234)."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY045825; AAK76499.1;  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synth.  
DR InterPro; IPR005856; Cys\_synth.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRPFAMs; TIGR01139; CysK; 1.  
DR TIGRPFAMs; TIGR01136; CysK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
KW Lyase.  
SQ SEQUENCE 321 AA; 33642 MW; 56531548939932B CRC64;

Query Match 84.2%; Score 1366.5; DB 10; Length 321;  
Best Local Similarity 83.1%; Pred. No. 6-4e-88;  
Matches 265; Conservative 30; Mismatches 23; Indels 1; Gaps 1;

QY 6 SGIADVTELLIGKTPLYINKLADGCVARVAAKLEIMPCSSVDORIGYMTADAEGL 65  
DB 3 SRIADVTELLIGNTPLV-LNNVAEGCVGVAAKLEIMPCSSVDORIGYMTADAEGL 61  
QY 66 ITPGKSVLEPTSGNTGIGLAFMAAARGKILITMPASVSEERRIILAFGAELVLTDP 125  
DB 62 IKRGSSVLEPTSGNTGIGLAFMAAARGKILITMPASVSEERRIILAFGAELVLTDP 121  
QY 126 KMKCAVQKAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGT 185  
DB 122 KMKCAVQKAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSGIGT 181  
QY 186 GGTITGAGKTYLKEQNPNIKLGVEPVESPVLSGKRGPHKIGIGAGFTPGVLEVNILDE 245  
DB 182 GGTITGAGKTYLKEQNPNIKLGVEPVESPVLSGKRGPHKIGIGAGFTPGVLEVNILDE 241  
QY 246 VVOISSDEAIEITAKLALKEGLFVGISGAAAAAFAOLAKRENNAGKLIIVAFPSFGERY 305  
DB 242 VVOISSDEAIEITAKLALKEGLFVGISGAAAAAFAOLAKRENNAGKLIIVAFPSFGERY 301

QY 306 LSVLPESVREARASMTFEP 324  
DB 302 LSVLPESVREARASMTFEP 320

RESULT 5

Q9FS26 PRELIMINARY; PRT; 386 AA.

AC Q9FS26  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Plastidic cysteine synthase 1.  
GN PCS-2-1.  
OS Solanum tuberosum (potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_Taxid=4113;  
RN [1]  
RP SEQUENCE FROM N.A.

RA STRAIN=Dansyaku; TISSUE=Sprout;  
RA Maruyama A., Ishizawa K.,  
RT "Plastidic Cysteine Synthase 1 (PCS-2-1) from Potato Sprout."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB029512; BAB20862.1;  
DR HSBP; P12674; 10AS.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synth.  
DR InterPro; IPR005856; Cys\_synth.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRPFAMs; TIGR01139; CysK; 1.  
DR TIGRPFAMs; TIGR01136; CysK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR PROSITE; PS00430; TONE\_DEPENDENT\_REC\_1; 1.  
SQ SEQUENCE 386 AA; 40519 MW; 82945B9D8640BC7 CRC64;

Query Match 81.1%; Score 1316; DB 10; Length 386;  
Best Local Similarity 78.4%; Pred. No. 2.9e-84;  
Matches 250; Conservative 37; Mismatches 32; Indels 0; Gaps 0;

QY 3 VERSGIADVTELLIGKTPLYINKLADGCVARVAAKLEIMPCSSVDORIGYMTADAE 62  
DB 4 IBSGNIADVTELLIGNTPLV-LNNVAEGCVGVAAKLEIMPCSSVDORIGYMTADAE 123  
QY 63 KGLITPGKSVLEPTSGNTGIGLAFMAAARGKILITMPASVSEERRIILAFGAELVLT 122  
DB 124 KGLITPGKSVLEPTSGNTGIGLAFMAAARGKILITMPASVSEERRIILAFGAELVLT 183  
QY 123 DPAKMGCAVQKAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSG 182  
DB 184 DPAKMGCAVQKAEELIAKTPNAVYILQGFENPANKVHYETTGPEIWKSGDKIDAFVSG 243  
QY 183 IGTGTTGAGKTYLKEQNPNIKLGVEPVESPVLSGKRGPHKIGIGAGFTPGVLEVNIL 242  
DB 244 IGTGTTGAGKTYLKEQNPNIKLGVEPVESPVLSGKRGPHKIGIGAGFTPGVLEVNIL 303  
QY 243 LDEVVOISSDEAIEITAKLALKEGLFVGISGAAAAAFAOLAKRENNAGKLIIVAFPSFG 302  
DB 304 LDEVVOISSDEAIEITAKLALKEGLFVGISGAAAAAFAOLAKRENNAGKLIIVAFPSFG 363  
QY 303 ERYLSSVLPESVREARASMTFEP 321

Db 364 ERYLSTLPOSIRECEK 382

## RESULT 6

09LJAO PRELIMINARY; PRT; 305 AA.  
 AC 09LJAO;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Cysteine synthase, O-acetylserine (thiol) lyase.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxId=3702;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu H., Tabata S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RX MEDLINE=20363099; PubMed=10907853;  
 RA Nakamura Y.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
 RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,  
 RT TAC and BAC clones."  
 RL DNA Res. 7:217-221(2000).  
 DR EMBL; AP000731; BAB01461.1; -.  
 DR HSP; P12674; I0AS.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; P:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR001216; Cys\_synthase\_BS.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthK.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; cysK; 1.  
 DR TIGRFAMs; TIGR01136; cysK; 1.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 KM Lyase.  
 SQ SEQUENCE 305 AA; 32338 MW; 130FA34FC47382C1 CRC64;

Query Match 79.9%; Score 1296; DB 10; Length 305;  
 Best Local Similarity 78.2%; Pred. No. 5.3e-83;  
 Matches 254; Conservative 28; Mismatches 23; Indels 20; Gaps 1;

QY 1 MAVERSGIAKQVTLIGKTPVYANKLADGCVAAFAAKLBIIEPCSSVKNRIGYSMTADA 60  
 DB 1 MAAAPKIAKQVTLIGKTPVYANKLADGCVAAFAAKLBIIEPCSSVKNRIGYSMTADA 60  
 QY 61 EERKGIITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 120  
 DB 61 EAKGILKRGESVLIETPSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 120  
 QY 121 LTDPKAGKGAIVQAEELIATKPNAYILQCFENPANKVHYETTGPEIMKSGDGIADAFV 180  
 DB 121 LTDPKAGKGAIVQAEELIATKPNAYILQCFENPANKVHYETTGPEIMKSGDGIADAFV 180  
 QY 181 SGIGTGTTGAGKYLKQNPNIKILGYEPVPSPVLSGKRGPKIQIGAGFTPGULVY 240  
 DB 181 SGIGTGTTGAGKYLKQNPNIKILGYEPVPSPVLSGKRGPKIQIGAGFTPGULVY 240  
 QY 241 NLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 300  
 DB 241 NLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 300  
 QY 221 DILDEVVOVSSSEIDNAKLAREBGLLVGISSGAAATPAIKAKRPENAGLIIVAFPS 280  
 DB 221 DILDEVVOVSSSEIDNAKLAREBGLLVGISSGAAATPAIKAKRPENAGLIIVAFPS 280  
 QY 301 FGERYLSVLPESVRRASMTFEP 325  
 DB 301 FGERYLSVLPESVRRASMTFEP 325

Db 281 FGERYLSVLPENAKREMTFEP 305

## RESULT 7

065747 PRELIMINARY; PRT; 266 AA.  
 AC 065747;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cysteine synthase, O-acetyl-L-serine (Thiol)-lyase (EC 4.2.99.8)  
 OS (Fragment).  
 OS Cicer arietinum (chickpea) (Garbanzo).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.  
 OX NCBI\_TaxId=3827;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Cv. Castellana; TISSUE=Etolated epicotyl;  
 RA Dopico B., Estebar R., Labrador B.;  
 RL "cDNA expressed in chickpea epicotyls."  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AU006024; CAA06819.1; -.  
 DR HSP; P12674; IFCU.  
 DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
 DR GO; GO:0016829; P:lyase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro; IPR001926; B6\_enzyme\_beta.  
 DR InterPro; IPR005859; Cys\_synthK.  
 DR InterPro; IPR005856; Cys\_synthK.  
 DR Pfam; PF00291; PALP; 1.  
 DR TIGRFAMs; TIGR01139; cysK; 1.  
 DR TIGRFAMs; TIGR01136; cysK; 1.  
 KM Lyase.  
 FT NON\_TER  
 SQ SEQUENCE 266 AA; 27911 MW; 1F253B511A356B CRC64;

Query Match 79.5%; Score 1290; DB 10; Length 266;  
 Best Local Similarity 94.7%; Pred. No. 1.2e-82;  
 Matches 253; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

QY 60 AERKGIITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 119  
 DB 1 AERKGIITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 119  
 QY 120 VLTDPKAGKGAIVQAEELIATKPNAYILQCFENPANKVHYETTGPEIMKSGDGIADAF 179  
 DB 120 VLTDPKAGKGAIVQAEELIATKPNAYILQCFENPANKVHYETTGPEIMKSGDGIADAF 179  
 QY 61 EERKGIITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 120  
 DB 61 EERKGIITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILLARGLV 120  
 QY 180 VSGIGTGTTGAGKYLKQNPNIKILGYEPVPSPVLSGKRGPKIQIGAGFTPGULVY 239  
 DB 121 VSGIGTGTTGAGKYLKQNPNIKILGYEPVPSPVLSGKRGPKIQIGAGFTPGULVY 180  
 QY 240 VNLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 299  
 DB 240 VNLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 240  
 QY 181 VNLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 240  
 DB 181 VNLIDDEVVOISDEAIETAKLALKEGLFVGISSGAAAAAFOIAKRPENAGLIIVAFPS 240  
 QY 300 SFGERYLSVLPESVRRASMTFEP 325  
 DB 241 SFGERYLSVLPESVRRASMTFEP 266

## RESULT 8

084SE4 PRELIMINARY; PRT; 409 AA.  
 AC 084SE4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative plastidic cysteine synthase 1.  
 DB P0020509.20.



RESULT 10  
081523 PRELIMINARY; PRT; 352 AA.  
ID 081523:  
AC 081523:  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DE Putative O-acetylserine (thiol) lyase precursor (EC 4.2.99.8).  
GN CRCY5-1A.  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OC NCBI\_TaxID=3055;  
RX [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94421691; PubMed=10491132;  
RA Ravita C.G., Barroso C., Vega J.M., Gotor C.;  
RT "Cysteine biosynthesis in Chlamydomonas reinhardtii. Molecular cloning  
and regulation of O-acetylserine(thiol) lyase.";  
RL Eur. J. Biochem. 264:848-853(1999).  
DR EMBL; AF078693; AAC2794.1; -.  
DR PIR; T07962; T07962.  
DR HSSP; P12674; 1OAS.  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; P:amino acid metabolism; IEA.  
DR GO; GO:0006530; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthase\_BS.  
DR InterPro; IPR005856; Cys\_synthase\_BS.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMs; TIGR01139; CYSK; 1.  
DR TIGRFAMs; TIGR01136; CYSK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR TRANSIT; TRANSIT; TRANSIT.  
FT CHAIN 32 352 POTENTIAL.  
FT SEQUENCE 352 AA; 36858 MW; 6A48C19231FEE79F CRC64;

Query Match 78.3%; Score 1270; DB 10; Length 352;  
Best Local Similarity 74.4%; Pred. No. 4.3e-81;  
Matches 241; Conservative 41; Mismatches 42; Indels 0; Gaps 0;  
DB 2 AVESGIAKDVTEIGKTPVLYNKLDGCVARVAKELMBPCSSVDRIGYSMTADAE 61  
DB 27 AAVGNKINATDTEIGKTPVLYNKLVAGTAKIAKLEIMBPCSSVDRIGYSMTADAE 86  
DB 62 EKGLITPKKSVLIPTSGNTGIGLAFAAAGVYLIITMPASMSLERIIILLAGAEVLT 121  
DB 87 KEGLTTPKTVLVEPTSGNTGIGLAFAAAGVYLIITMPASMSLERIIILLAGAEVLT 146  
DB 122 TDPKMGKGAOKAEELIATPNAYILLOQFENPANKVHYETTGELIKWSDGKIDAFVS 181  
DB 147 TDPKMGKGAOKAEELIATPDAPMLDQFQNPNNPKHYETTGELIKWSDGKIDAFVS 206  
DB 182 GIGTGGITGAGKYLKEONPNIKLIGVEPVSPVLSGGKPGPHKIQIGIAGFIPGVLEVN 241  
DB 207 GVGITGGITGAGKYLKEONPNIKLIGVEPVSPVLSGGKPGPHKIQIGIAGFIPGVLEVN 266  
DB 242 LDEVOVSSDEALETATLAKLKEGLFVGISSGAAAPOLAKPENAGKLIYAVPSPF 301  
DB 267 LISEVVOVSSDAILMARIALREGLMGISSGAAVQALIVASRPNKGLVIVVLPF 326  
DB 302 GERYLSSVLFESVRRASMTPEP 325  
DB 327 GERYLSSVLFQOLRDEASKMTPEP 350

AC 09F825;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DE Plastidic cysteine synthase 2.  
GN PCS-2-2.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OC NCBI\_TaxID=4113;  
RX [1]  
RP SEQUENCE FROM N.A.  
RP STRAIN=Dansyaku; TISSE=Sprout;  
RA Maruyama A., Ishizawa K.;  
RT "Plastidic Cysteine Synthase (PCS-2-2) from Potato Sprout.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB029513; BAB20863.1; -.  
DR HSSP; P12674; 1OAS.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004124; F:cysteine synthase activity; IEA.  
DR GO; GO:0016829; P:lyase activity; IEA.  
DR GO; GO:0004872; P:receptor activity; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006520; P:amino acid metabolism; IEA.  
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.  
DR InterPro; IPR001926; B6\_enzyme\_beta.  
DR InterPro; IPR001216; Cys\_synthase\_BS.  
DR InterPro; IPR005859; Cys\_synthase\_BS.  
DR InterPro; IPR005856; Cys\_synthase\_BS.  
DR InterPro; IPR00531; TonB\_boxC.  
DR Pfam; PF00291; PALP; 1.  
DR TIGRFAMs; TIGR01139; CYSK; 1.  
DR TIGRFAMs; TIGR01136; CYSK; 1.  
DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; 1.  
DR SEQUENCE 376 AA; 39960 MW; 47FDD6EE2C6A75C CRC64;

Query Match 76.8%; Score 1246; DB 10; Length 376;  
Best Local Similarity 75.2%; Pred. No. 2.3e-79;  
Matches 240; Conservative 37; Mismatches 32; Indels 10; Gaps 1;  
DB 3 VBSGIAKDVTEIGKTPVLYNKLDGCVARVAKELMBPCSSVDRIGYSMTADAE 62  
DB 64 IEGTINMEDVTOLIGNTPMYLIATKGCYANIAKLEIMBPCSSVDRIGYSMTADAE 123  
DB 63 KGLITPKKSVLIPTSGNTGIGLAFAAAGVYLIITMPASMSLERIIILLAGAEVLT 122  
DB 124 KGLITPKKSVLIPTSGNTGIGLAFAAAGVYLIITMPASMSLERIIILLAGAEVLT 183  
DB 123 DPAKMGKGAOKAEELIATPNAYILLOQFENPANKVHYETTGELIKWSDGKIDAFVS 182  
DB 184 DPAKMGKGAOKAEELIATPDAPMLDQFQNPNNPKHYETTGELIKWSDGKIDAFVS 243  
DB 183 IGTGGITGAGKYLKEONPNIKLIGVEPVSPVLSGGKPGPHKIQIGIAGFIPGVLEVN 242  
DB 244 IGTGGITGAGKYLKEONPNIKLIGVEPVSPVLSGGKPGPHKIQIGIAGFIPGVLEVN 293  
DB 243 LDEVOVSSDEALETATLAKLKEGLFVGISSGAAAPOLAKPENAGKLIYAVPSPF 302  
DB 294 MDEVIRISSDEALETATLAKLKEGLFVGISSGAAAPOLAKPENAGKLIYAVPSPF 353  
DB 303 ERYLSSVLFESVRRASMTPEP 321  
DB 354 ERYLSSVLFQOLRDEASKMTPEP 372

RESULT 11  
09F825 PRELIMINARY; PRT; 376 AA.  
ID 09F825:  
AC 043726;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

RESULT 12  
043726 PRELIMINARY; PRT; 324 AA.  
ID 043726:  
AC 043726;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)



DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)  
 DE Cysteine synthase (EC 4.2.99.8).  
 GN CYS ACS 1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsi.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=flower;  
 RA MEDLINE=95334513; PubMed=761018;  
 RT Hesse H., Altmann T.,  
 RT "Molecular cloning of a cysteine synthase cDNA from Arabidopsis  
 thaliana.".  
 RL Plant Physiol. 108:851-852(1995).  
 DR EMBL: X81697; CA57343.1; -.  
 DR PIR: S49586; S49586.  
 DR HSSP: P12674; 1OAS.  
 DR GO: GO:0004124; P:cysteine synthase activity; IEA.  
 DR GO: GO:0016829; P:lyase activity; IEA.  
 DR GO: GO:0006520; P:amino acid metabolism; IEA.  
 DR GO: GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro: IPR001926; B6 enzyme beta.  
 DR InterPro: IPR005859; Cys\_synthK.  
 DR InterPro: IPR005856; Cys\_synthKM.  
 DR Pfam: PF00291; PALP; 1.  
 DR TIGRFAMs: TIGR01139; cysK; 1.  
 DR TIGRFAMs: TIGR01136; cysKM; 1.  
 DR Lyase.  
 SQ SEQUENCE 324 AA; 34605 MW; 6F43A718C1B3E1BD CRC64;  
 Query Match 75.7%; Score 1229; DB 10; Length 324;  
 Best Local Similarity 75.9%; Pred. No. 2.9e-78;  
 Matches 246; Conservative 34; Mismatches 38; Indels 6; Gaps 4;  
 QY 6 SGIAKVTLLIGTPIYLYNKLKLA-DGCVARVAANKLEMEPC--SYKDRIGYSMIDAE 62  
 DB 3 SRKAVTELLIGTPIYLYNKLKLA-DGCVARVAANKLEMEPCSVKDRIGYSMIDAE 62  
 QY 63 KGLITPKSVLYLEPTSGNTG-IGLAPMAAAGKGLIITMPASVSLERRITLAFGLAVL 121  
 DB 63 KGLIKGSEVLYLEPTSGNTGELLAFPAACKGYKLIITMPASVSLERRITLAFGLAVL 122  
 QY 122 TPDAKCKKGVCAABRIILAKTPNAYLIOGFENPANKVHETTGPELTKGSDKIDAFVS 181  
 DB 123 TPDAKCKKGVCAABRIILAKTPNAYLIOGFENPANKVHETTGPELTKGSDKIDAFVS 182  
 QY 182 GIGTGGITTAGKYLKRONENIKLIGVPEVPSVLSGKXPKPKIOGIGAGTIPGLAVN 241  
 DB 183 GIGTGGITTAGKYLKRONENIKLIGVPEVPSVLSGKXPKPKIOGIGAGTIPGLAVN 242  
 QY 242 LADVAVQISSDAIRKAKLALKEGLFVGISSGAAAAAQAQIKRPNMGKLIIVAPSP 301  
 DB 243 LADVAVQISSDAIRKAKLALKEGLFVGISSGAAAAAQAQIKRPNMGKLIIVAPSP 300  
 QY 302 GERYTSSVLFESVREARESWTFEP 325  
 DB 301 GERYTSSVLFESVREARESWTFEP 324  
 RESULT 13  
 Q9W313 PRELIMINARY; PRT; 282 AA.  
 AC Q9W313;  
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)  
 DE Cysteine synthase, 5'-partial (fragment).  
 GN OsUNAB0069E14.1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=45530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gamberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,  
 RA Teitelin T., Riggs F., Heico J., Zeman V., Blunt S., Pai G.,  
 RA Vanaken S.E., Utterback T.R., Feldlyum T.V., Kalb B., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSUNAB0069E14 genomic sequence.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC091811; AAL58961.1; -.  
 DR Gramene: Q9W313; -.  
 DR GO: GO:0004124; P:cysteine synthase activity; IEA.  
 DR GO: GO:0016829; P:lyase activity; IEA.  
 DR GO: GO:0006520; P:amino acid metabolism; IEA.  
 DR GO: GO:0006535; P:cysteine biosynthesis from serine; IEA.  
 DR InterPro: IPR001926; B6 enzyme beta.  
 DR InterPro: IPR005859; Cys\_synthK.  
 DR InterPro: IPR005856; Cys\_synthKM.  
 DR Pfam: PF00291; PALP; 1.  
 DR TIGRFAMs: TIGR01139; cysK; 1.  
 DR TIGRFAMs: TIGR01136; cysKM; 1.  
 DR NON TER 1  
 SQ SEQUENCE 282 AA; 29892 MW; 3503D2F3CA80FBF5 CRC64;  
 Query Match 75.6%; Score 1227; DB 10; Length 282;  
 Best Local Similarity 82.7%; Pred. No. 3.3e-78;  
 Matches 235; Conservative 24; Mismatches 23; Indels 2; Gaps 1;  
 QY 42 MEPCSSVDKRIQYSMIDAEKGLITTPGKSVLYLEPTSGNTGIGLAPMAAARGKYLITMP 101  
 DB 1 MEPCSSVDKRIQYSMIDAEKGLITTPGKSVLYLEPTSGNTGIGLAPMAAARGKYLITMP 60  
 QY 102 ASMSLEERRIILAFGLAVLTPDAKCKKGVCAABRIILAKTPNAYLIOGFENPANKVH 161  
 DB 61 ASMSLEERRIILAFGLAVLTPDAKCKKGVCAABRIILAKTPNAYLIOGFENPANKVH 120  
 QY 162 EPTGPEIMKGSQKIDAFVSIGTGTITGAGKYKRONENIKLIGVPEVPSVLSGKRP 221  
 DB 121 EPTGPEIMKGSQKIDAFVSIGTGTITGAGKYKRONENIKLIGVPEVPSVLSGKRP 180  
 QY 222 GPHKIQIGAGFIPGLAVNLDVAVQISSDAIRKAKLALKEGLFVGISSGAAAAAF 281  
 DB 181 GPHKIQIGAGFIPGLAVNLDVAVQISSDAIRKAKLALKEGLFVGISSGAAAAAF 240  
 QY 282 QIAKRPENAGKLIIVAPSPGERTYLSVLFESVREARESWTFEP 325  
 DB 241 QIAKRPENAGKLIIVAPSPGERTYLSVLFESVREARESWTFEP 282  
 RESULT 14  
 Q9FSF5 PRELIMINARY; PRT; 324 AA.  
 AC Q9FSF5;  
 DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Cysteine synthase.  
 GN GSI.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shen W.H.;  
 RT "Nicotiana tabacum cDNA encoding cysteine synthase (O-acetylserine  
 sulfinylase).";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.





Mon May 24 08:18:53 2004

us-09-931-457a-31.rtf

Page 2

```

RESULT 2
US-09-543-681A-7733
Sequence 7733, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BEETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7733
LENGTH: 335
TYPE: PRF
ORGANISM: Proteus mirabilis
US-09-543-681A-7733

```

Query Match	882;	Score	882;	DB 4;	length	335;			
Best Local Similarity	56.6%;	Pred. No.	1.6e-81;						
Matches	181;	Conservative	51;	Mismatches	76;	Indels	12;	Gaps	5;

```

Qy      VSSSGAKVNTLEIGTPTVYLNKADGCVARVAKAEIMEPCSSVDORIGYSLMDAER 62
        :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      17 IRMSKIFEDNSQITGHTPLVRLKHGNG---NLAKVESRNPSSVCKRIGANIMIDAEK 73

Qy      63 KGLITPGKSVLLEPTSGNTGIGLAFMAAARGKYLITTPAMSTERRIILAEAEIYLT 122
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      74 KGLIKPGVE-LVEPTSGNTGIALAAVAAARGKYLITPDMTSEIRRLKALGANVLT 132
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      123 DPAKMGKGAVOKAEELIATTPNAY-ILQGFENPAMPKVHYETTPBEIWKSGDKIDAEVS 181
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      133 EGAKMGKGAIDKRNELRSDSPSRYLLIQGFSPNAPBEIHEKTGTGEIMNDTDGADVAVIA 192
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      182 GIGTGTITGAGKYLKE-QNPINKLIGVPPVSPULSGS-----KGGPHKIGTIGAGFI 234
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      193 GVGGTGTTIGIKRYLKTQGAQKATWVAEPKOSPITSQLAAGEBIRKEGPHKIGTIGAGFI 252
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      235 PGVLEVNLLDEVAQISDEAIEPAKLLAKBGLFVGISGAAAAAARQIAAPENNAKGLI 294
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      253 FENLEISLDRVITQFNBEAPEATARDVMTKEGILAGISGAALAAAVLAKEPEFANKKI 312
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      295 VAVPSPGGEYLSSTVPESSV 314
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      313 VVLPSPGGEYLSSTVPEADI 332
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
US-09-328-352-5450
; Sequence 5450, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Galy L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5450
; LENGTH: 335
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5450

Query Match          53.3%; Score 865; DB 4; Length 335;
Best Local Similarity 56.3%; Pred. No. 8,7e-80;
Matches 119; Conservative 52; Mismatches 75; Indels 12; Gaps 6;

```

Dh 17 GAAVSNNAEAGTFLVRINRLIK-  
TATATLAVAYESNNPFAVYKRI  
GALLIADAEKRG 75

Qy 65 LITPGASVLIERTSGNTGIGLA  
FMAAAGYGLIITMPSMSIERIIL  
LAFGAELVITDP 124

Dh 76 ALKEGHNI-VEETSGNTGIALA  
FVAAGYITLTMPASMSIERKXV  
KOLGALVITEP 134

Qy 125 AKGMRGAVOKAESIIAKT  
ENAYIL-QQFENANPKVAYETTG  
PEIWKSGDGLDAFVSGI 163

Dh 135 AKGMRGAVBEAVRLATE  
OPEVYFTLQQGFENANPQI  
HVDITGPBIWATGQAVDILVAGV 194

Qy 184 GGGGTITGAGKYLKE-QNP  
NIKLGVEFVSPVLSGGK-----  
PGPHKIQIGIGAGEIPG 236

Dh 195 GGGGTITGTSKRFQONKPL  
SVAVEPABESPIITQTKNGENIT  
PAHHKIQIGIGANIPK 254

Qy 237 VLEGNILDEVOVLISSDE  
ALETAKLIALKGLFVGISGAAAA  
A-PQIAKRPENNGKILVA 296

Dh 255 NMLDIDLVDEVLFPVSS  
SEBAIQMARCATQBGILVIGISG  
AALAAVAKIARPENNGKITV 314

Qy 297 VFPSPGERTLSSVLPFSV  
314

Dh 315 ILPDGERTLSSVLPBGL 332

```

RESULT 4
US-09-252-991A-31491
; Sequence 31491, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,768
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 3142
; SEQ ID NO 31491
; LENGTH: 748
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31491

```

Query Match	52.7%	Score 866;	DB 4;	Length 748;
Beet Local Similarity	56.5%	Pred. No. 2,5e-78;		
Matches	182;	Conservative	46;	Mismatches 84; Indels 10; Gaps 6

  

QY	6	SGIADVTELLIGKTPLVYINFLKADGCVARVAKILMEBCSSVKORIGYSIMADABEKL	65
DB	426	SRIFADNQSIGNTFVLQINRIAPRGV--IIAKTIGRRPGYSVKCRIGANMTWDAESGR	484
QY	66	ITPGKSVLIEPSSNTGIGLAFMAARGYKILTMPASNSLERRIILILAFGALVLTGDA	125
DB	485	LKSGMT-LVEPSSNTGIGLAFVAAARGKILITMPASNSLERRKYLKALGALVLTGDA	543
QY	126	KGMKGAOVAKAEIILAKTPNAVYL--QGFENPAPKPHYRYTGEIIMKSGSDKIDAFVSGIG	184
DB	544	KMKRAITQKAEELVAGDPEKTPMPOFDNPAPFAIHEKTIIGEININMDGAVDLVLSGVG	603
QY	185	TGGTITGAGKYLYKE-QNPNIKILGYEPVESPVLS--GG--KPGPHKIQSIGAGGTPGV	237
DB	604	TGGTLTGVSRYIKNTRGKPIILAAVEPPLTSPVISTGLAGEEYKAPPHKIQSIGAGGVFAN	663
QY	238	LEVNILDEVQSSSREAIETAKLILKEGLFVIGISGGAATAAFAOIAKRPENNAKGIIVAV	297
DB	664	LDLSTVDRPKIGDDEAKNNALRIIMOBSLITGIGSGAAMAAVRLAIEPPNNQGIIVVI	723
QY	298	PPSGGRYLSVLPFESVRRAE 319	
DB	724	LPDSGRYLSKTFGLFSRGE 745	

RESULT 5

US-09-489-039A-9166  
 Sequence 9166, Application US/09489039A  
 Patent No. 6610836  
 GENERAL INFORMATION:  
 APPLICANT: Gary Breton et. al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709 2004001  
 CURRENT APPLICATION NUMBER: US/09/489, 039A  
 CURRENT FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: US 60/117,747  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 14342  
 SEQ ID NO 9166  
 LENGTH: 326  
 TYPE: PRN  
 ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-9166

Query Match Score 853.5; DB 4; Length 326;  
 Best Local Similarity 54.2%; Pred. No. 1.2e-78;  
 Matches 175; Conservative 56; Mismatches 79; Indels 13; Gaps 5;

QY 6 SGIAADVTELLIGTPTLVYLNKLADGCVARVAATLMEPCSSVKDRIGYSMTDAEKG 65  
 DB 5 SKIFEDNSLTIGTPTLVYLNKLADGCVARVAATLMEPCSSVKDRIGYSMTDAEKG 61  
 QY 66 ITPGVSYLEPTSGNIGLAAARAGKLIITPMSLERRIILAFGAEVLTDP 125  
 DB 62 ITPGVSYLEPTSGNIGLAAARAGKLIITPMSLERRIILAFGAEVLTDP 120  
 QY 126 KMKGAQVAKAEIILAKTPNVA-110QFENPANKVYETTPETIKSGDKIDAFVSGIG 184  
 DB 121 KMKGAQVAKAEIILAKTPNVA-110QFENPANKVYETTPETIKSGDKIDAFVSGIG 180  
 QY 185 TCGTTGAGKTYLKEONPNKLI--GYEPVPSVLSG-----KGPKEIOGIGAGPTPG 236  
 DB 181 TCGTTGAGKTYLKEONPNKLI--GYEPVPSVLSG-----KGPKEIOGIGAGPTPG 240  
 QY 237 VEVNVLIDVVQISSDEAIEFATLAKLKEGLFVGISGGAIAAARFQIARPNAGKLIYA 296  
 DB 241 NIDLVKRVIGITTHEALSTARLMEHGIILGIISSGAAVAAALKLDEAFTHKNTIV 300  
 QY 297 VFPFGERYLSVLPESVEREAP 319  
 DB 301 ILPSSGERYLSVLPESVEREAP 323

RESULT 6

US-09-107-532A-3925  
 Sequence 3925, Application US/09107532A  
 Patent No. 6583275  
 GENERAL INFORMATION:  
 APPLICANT: Lynn A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: PC  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA: US/09/107 532A

FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/065,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 3925:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 308 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium

FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (8) LOCATION 1...308  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3925:  
 US-09-107-532A-3925

Query Match Score 813; DB 4; Length 308;  
 Best Local Similarity 54.8%; Pred. No. 1.5e-74;  
 Matches 171; Conservative 50; Mismatches 85; Indels 6; Gaps 5;

QY 5 RSGIAADVTELLIGTPTLVYLNKLADGCVARVAATLMEPCSSVKDRIGYSMTDAEKG 64  
 DB 1 RDAFNSVTELLIGTPTLVYLNKLADGCVARVAATLMEPCSSVKDRIGYSMTDAEKG 60  
 QY 65 LITPKSVLIEPTSGNIGLAAARAGKLIITPMSLERRIILAFGAEVLTDP 124  
 DB 61 LITPKSVLIEPTSGNIGLAAARAGKLIITPMSLERRIILAFGAEVLTDP 119  
 QY 125 AKMGGAQVAKAEIILAKTPNVA-110QFENPANKVYETTPETIKSGDKIDAFVSGIG 182  
 DB 120 AKMGGAQVAKAEIILAKTPNVA-110QFENPANKVYETTPETIKSGDKIDAFVSGIG 177  
 QY 183 TCGTTGAGKTYLKEONPNKLI--GYEPVPSVLSG-----KGPKEIOGIGAGPTPG 242  
 DB 178 TCGTTGAGKTYLKEONPNKLI--GYEPVPSVLSG-----KGPKEIOGIGAGPTPG 237  
 QY 243 LDEVNVLIDVVQISSDEAIEFATLAKLKEGLFVGISGGAIAAARFQIARPNAGKLIYA 302  
 DB 238 YDKVLSISGDBAMETAEVGERKRIILVIGISGGAIAAARFQIARPNAGKLIYA 296  
 QY 303 ERYLSVLPESV 314  
 DB 297 ERYLSVLPESV 308

RESULT 7

US-09-114-001C-3330  
 Sequence 3330, Application US/09134001C  
 Patent No. 6380370  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: GTC-007  
 CURRENT APPLICATION NUMBER: US/09/134,001C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: US 60/055,779  
 PRIOR FILING DATE: 1997-08-14  
 NUMBER OF SEQ ID NOS: 5674

Mon May 24 08:18:53 2004

US-09-931-457a-31.rat

Page 4

SEQ ID NO 3330  
LENGTH: 319  
TYPE: PRN  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3330

Query Match 49.4%; Score 801; DB 4; Length 319;  
Best Local Similarity 52.6%; Pred. No. 2,7e-73;  
Matches 159; Conservative 60; Mismatches 77; Indels 6; Gaps 5;

QY 12 VTELLIGTPTVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLITGKS 71  
DB 18 VTQITGTPVVKLANVDDADADIVKLEYNPGQSVKRIALAMIEKAREGKIKRGDT 77  
QY 72 VLIEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKGKA 131  
DB 78 I-VEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKGKA 136  
QY 132 VOKAEELIAKTPNAYI-LQOFENPANKVHYETTPRIKMSDCK-IDAFVSGIGNGGTI 189  
DB 137 IKKAKEL--KEHGFPEPQOFENPANKVHYETTPRIKMSDCK-IDAFVSGIGNGGTI 194  
QY 190 TGAGKYLKEQNPNKILGVEFVESPVLSGGKPGPHKIQIGAGFTPGVLBNLDEVYQI 249  
DB 195 SGVGKVLKKEYPNVEIYVIEPASPVLSSGEPGPHKIQIGAGFTPGVLBNLDEVYQI 254  
QY 250 SSDAIEIAKLLAKELFVIGISSGAAAAAFOIAKPEENAKILVAVPSPGGRYLSV 309  
DB 255 GNDPAMMARVAEKEGILAGISSGAAIYALIOAKE-LGKKTIVTVLPSNGERYLSTP 313  
QY 310 LF 311  
DB 314 LV 315

RESULT 8  
US-09-724-623-72  
Sequence 72, Application US/09724623  
Patent No. 6476209  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Lubbers, Mark W  
TITLE OF INVENTION: Polynucleotides, materials incorporating  
FILE REFERENCE: 1048U1  
CURRENT APPLICATION NUMBER: US/09/724,623  
NUMBER OF SEQ ID NOS: 124  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 72  
LENGTH: 309  
TYPE: PRN  
ORGANISM: Lactobacillus rhamnosus  
US-09-724-623-72

Query Match 46.9%; Score 761; DB 4; Length 309;  
Best Local Similarity 51.5%; Pred. No. 3.1e-69;  
Matches 157; Conservative 56; Mismatches 88; Indels 4; Gaps 4;

QY 9 AKDVTIELIGTPTVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLITP 68  
DB 5 ADNTIGTIGTPTVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLITP 64  
QY 69 GKSVALIEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKG 128  
DB 65 G-GTIVEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKG 123  
QY 129 KGAOKAEELIAKTPNAYI-LQOFENPANKVHYETTPRIKMSDCK-IDAFVSGIGNGGTI 187  
DB 124 PGALIKKA-EALSKENGITPLMFOFONPANDVHERITGELIRSDGIGPDAFVAGVGTGG 182  
QY 188 TITGAGKYLKEQNPNKILGVEFVESPVLSGGKPGPHKIQIGAGFTPGVLBNLDEVY 247

DB 183 TITGVRALRKINPVDQIVALEPABSPMLKEHGGHKKIQIGAGFTPVLDTNLXYDII 242  
QY 248 QISSDAIEIAKLLAKELFVIGISSGAAAAAFOIAKPEENAKILVAVPSPGGRYLS 307  
DB 243 EYTSDAIDMAHVSHERBGLFPOISGANTIPGALIKK-LGKGSVAATVAPDNGERYLS 301  
QY 308 SVLFE 312  
DB 302 TDLEFX 306

RESULT 9  
US-09-543-681A-7720  
Sequence 7720, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB  
FILE REFERENCE: 2709,1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
PRIOR FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7720  
LENGTH: 303  
TYPE: PRN  
ORGANISM: Proteus mirabilis  
US-09-543-681A-7720

Query Match 40.1%; Score 650.5; DB 4; Length 303;  
Best Local Similarity 45.4%; Pred. No. 5.5e-58;  
Matches 138; Conservative 52; Mismatches 103; Indels 11; Gaps 4;

QY 8 IANDVTIELIGTPTVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLIT 67  
DB 10 IVAGLEQFIGNTPVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLIT 69  
QY 68 PKGSVALIEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKG 127  
DB 70 PG-DTIEPTSGNTGIGLAFAAARGYLLITPMSVSLERRIILAFGAELVITDPKAKG 128  
QY 128 MKGAVOKAEELIAKTPNAYI-LQOFENPANKVHYETTPRIKMSDCK-IDAFVSGIGNGGTI 187  
DB 129 MEGARDIAGE-MERKEGKVLIDQFNNDPRAHFTSTGEIIMQOTGRITHEFVSMGTG 187  
QY 188 TITGAGKYLKEQNPNKILGVEFVESPVLSGGKPGPHKIQIGAGFTPGVLBNLDEVY 247  
DB 189 TITGVSGLYLTQSTVYLNKLDGCVARVAALKEMLPEPCSSVQDRIGYSMTDAEKEGLIT 240  
QY 248 QISSDAIEIAKLLAKELFVIGISSGAAAAAFOIAKPEENAKILVAVPSPGGRYLS 307  
DB 241 DMSGTBAEQMRLASBGIFGVSSGAAVAGALRYAK--ENPGAVTVAIICDRGORYLS 298  
QY 308 SVLFE 311  
DB 299 TGVF 302

RESULT 10  
US-09-489-039A-9175  
Sequence 9175, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709,2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747

Mon May 24 08:18:53 2004

us-09-931-457a-31.ra1

Page 5

PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 9175  
LENGTH: 306  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9175

Query Match  
Best Local Similarity 38.7%; Score 628.5; DB 4; Length 306;  
Matches 135; Conservative 56; Mismatches 102; Indels 11; Gaps 4;

QY 8 IAKDVTIELIGTPLYLNKADGCVARVAAKIEMEPSCSVKDRIGYKMLDAEKGILIT 67  
DB 3 IVNTLEQITGNTPLYLNKADGCVARVAAKIEMEPSCSVKDRIGYKMLDAEKGILIT 62  
QY 68 PEKSVLIEPTSGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANG 127  
DB 63 PG-DVLIETSGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANG 121  
QY 128 MKGAVOKAEELIAKTPENAVYIIQOFENPANKVHYETTGEBIWKSGDKIDAFVSGIGTGG 187  
DB 122 MEGARDIALE-MAORGEKGLDQFNNPDPVAVHYTTTGTGEIWOCTAGRIHFWSSMGITG 180  
QY 188 TTTGAKYIKKQNPNIKIGVPEVPSVLSGKPGPHKIOGIGAGPIPGVLENTLDEVY 247  
DB 181 TTTGVSRLFRESKPTVIGLQPEBSSIPGIRPMP-----ATMPEITFNASLVDTVL 233  
QY 248 QISDEAIEFTAKILAKELGFVGISSGAAAAAFOIAKPPENAKLIYAVFSPFGERYLS 307  
DB 234 DIHQDAENTMQLAVREGIFGVSSGAVAGALRIAR--ENPGAVVAIVCDRGRYLS 291  
QY 308 SVLF 311  
DB 292 TGVF 295

RESULT 11  
US-09-252-991A-29566  
Sequence 29566; Application US/09252991A  
PATENT No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 29566  
LENGTH: 382  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29566

Query Match  
Best Local Similarity 41.9%; Score 598.5; DB 4; Length 382;  
Matches 126; Conservative 60; Mismatches 104; Indels 11; Gaps 4;  
QY 12 VTEIIGTPLYLNKADGCVARVAAKIEMEPSCSVKDRIGYKMLDAEKGILITPKS 71  
DB 91 IACVGTPTPLVRLQRLRGESNTLVLVEGNNPAGSVKDRPALSMITRALRGDIRPG-D 149  
QY 72 VLIETSGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANGKGA 131  
DB 150 TLIETSGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANGKGA 209  
QY 132 VQAEELIATPNAVYIIQOFENPANKVHYETTGEBIWKSGDKIDAFVSGIGTGGITG 191

DB 210 RDLADK-LQREGRKVLDDQFANGDNPNRAHNSCPETIWOOTGSSITHFVSSMGITGIMG 268  
QY 192 AKYIKKQNPNIKIGVPEVPSVLSGKPGPHKIOGIGAGPIPGVLENTLDEVYQISS 251  
DB 269 VSRVYIKKQNPNAVYIIQOFENPANKVHYETTGEBIWKSGDKIDAFVSGIGTGGITG 321  
QY 252 DEATETAKILAKELGFVGISSGAAAAAFOIAKPPENAKLIYAVFSPFGERYLSVLF 311  
DB 322 DEADIDIRLAREREGIFGVSSGAVAGALRIAR--ENPDVAVIICDRGRYLSGTV 379  
QY 312 E 312  
DB 380 D 380

RESULT 12  
US-09-328-352-8179  
Sequence 8179; Application US/09328352  
PATENT No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-039A  
CURRENT APPLICATION NUMBER: US/09/328,352  
PRIOR FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 8179  
LENGTH: 307  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-8179

Query Match  
Best Local Similarity 45.9%; Score 594; DB 4; Length 307;  
Matches 136; Conservative 49; Mismatches 97; Indels 14; Gaps 5;  
QY 16 IGTPTVLYLNKADGCVARVAAKIEMEPSCSVKDRIGYKMLDAEKGILITPKSVTIE 75  
DB 22 VGTPTVLYLNKADGCVARVAAKIEMEPSCSVKDRIGYKMLDAEKGILITPKSVTIE 80  
QY 76 PISGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANGKGAVOKA 135  
DB 81 AISGNTGIGLAFMAAARGYKLIITPASMILERRIILAFGAELVITDPANGKGAVOKA 137  
QY 136 BEILAKTPENAVYIIQOFENPANKVHYETTGEBIWKSGDKIDAFVSGIGTGGITGAGY 195  
DB 138 LQMOLE-GGLVINGPAGNDVNAVHYTTTGTGEIWOCTAGRIHFWSSMGITGIVSKY 196  
QY 196 LKQNPNIKIGVPEVPSVLSGKPGPHKIOGIGAGPIPGVLENTLDEVYQISSDAI 255  
DB 197 LKQNPNDIIGLQPSBSNINAGIRPQB-----VLPTEPERKVDQINDIQIARE 249  
QY 256 ETKAKILAKELGFVGISSGAAAAAFOIAKPPENAKLIYAVFSPFGERYLSVLF 311  
DB 250 KTRRLAREREGIFGVSSGAVAGALRIAR--ENPDVAVIICDRGRYLSGTV 303

RESULT 13  
US-09-540-236-3342  
Sequence 3342; Application US/09540236  
PATENT No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
PRIOR FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 3342  
LENGTH: 310  
TYPE: PRT

Mon May 24 08:18:53 2004

us-09-931-457a-31.ra1

Page 6

ORGANISM: M. catarrhalis  
US-09-540-236-3342

Query Match 36.5%; Score 593; DB 4; Length 310;

Best Local Similarity 42.5%; Pred. No. 4, 2e-52;  
Matches 135; Conservative 56; Mismatches 105; Indels 20; Gaps 6;

QY 1 MAVERSGIARD-----VTELLGKPLVYANKLADCCVAVAAKELMEPCSSVDRIG 53  
DB 2 MIYETITMTD FHIQTTLTSDCVNTPVYKNTLAKSOATLLAKEGNNPAGSVDRPA 61  
QY 54 YSMIADAEKELTPGSGVLEPTSGNTGIGLAFMAARGKLIITPASKSLERRITLL 113  
DB 62 FNNIYOEKQKQIRPG-DTLEBATSNTGIALANVAMRGYOMTILMPASTGERDAMT 120  
QY 114 AFGAEVLVLPDPAKCKGAVOKAEELIAKTPNAVILQQFENPANKVHYETTGPEIKGSD 173  
DB 121 AYGATLE--EVADGMEARPMALQMQA-DEGGIVLDQFNNITDNNKHAHYLTGPELMQOTE 177  
QY 174 GKIDAPVSGIGTGTGAKKYLKEONPNIKLIGVEEVESPVLSGKRGPHKIQIGIAGF 233  
DB 178 GKITHPVSWGTGTGVSQYLEONPALQIIGLOPDEGSSINGIRRP-----AYE 230  
QY 234 IPGVLEVNLDEVVQISSDENIETAKLALKEGLFVGISGAAAAAFQIAKPEENAKGL 233  
DB 231 LPGIFDASIVDTIMDVNQNAERVMRLAREEGIFGVSSGGAAMAHQIAKTPDA--V 288  
QY 294 IVAVPSFGERYLSSVLP 311  
DB 289 IAPVCDRGDRYLTGLF 306

RESULT 14

US-09-789-300A-4  
Sequence 4, Application US/09789300A

Patent No. 6458576  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel  
APPLICANT: Rudolph-Owen, Laura A.  
TITLE OF INVENTION: Dependent Enzyme Family Member and Uses Therefore  
FILE REFERENCE: 35860/208926  
CURRENT APPLICATION NUMBER: US/09/789,300A  
PRIOR FILING DATE: 2001-02-20  
PRIOR APPLICATION NUMBER: US 60/183,208  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Pyridoxal-Phosphate Dependent Enzyme Family Domain  
OTHER INFORMATION: Sequence  
US-09-789-300A-4

Query Match 36.0%; Score 584; DB 4; Length 378;

Best Local Similarity 43.0%; Pred. No. 4, 7e-51;  
Matches 161; Conservative 45; Mismatches 80; Indels 88; Gaps 17;

QY 12 VTELLGKPLVYANKLADCCVAVAAKELMEPCSSVDR-IGYSMIADAE 61  
DB 1 VTELLGKPLVYANKLADCCVAVAAKELMEPCSSVDR-IGYSMIADAE 60  
QY 62 ---EKGLTPG-----KSVLEPTSGNTGIGLAFMAARGKLIITPASKSLERRIT 111  
DB 61 KLGKGGIVPGTGVESKTYIIIEPTSGNTGIALALMAALLGKTIIVPATDTSREKTAQ 120  
QY 112 LIAAGARIVLTPDAK---KGAIVOKAEELIATP--NAVILQQ-----FENPANKV--HY 161  
DB 121 FRAUGAEIVVVPVAGGSDADADIAKASAEKPEENAVILNQAPGFDPNAPETAGQ 180

QY 162 ETTGPEIKMGSDPK-----IDAFVSGIGTGTGAKKYLKEONPN-----IKL 205  
DB 181 KTIIGPEIMQLGKELISGRLPDAVAVAPVGGGTTTGARVYKELNPDGKIDVLSIPVY 240  
QY 206 IGVEPESPVLSG-----GKPGP-----HKIIGIAG----- 232  
DB 241 IGVEPESAVLSGSKATVTLGKPLGGRDSKYLQDEPVLPETKSGIGIGLVPRVG 300  
QY 233 -FLPGVLEVNL-----LDEVQISSDEALETAKLALKEGLFVGISGAAAAAFQIA 284  
DB 301 EFPVPLDELDRROGIDEVTVTDEALEARLAREEGILVGPSSGAVALALAKLAE 360  
QY 285 -KPEENAGLIVAV 297  
DB 361 GKXPLNKGTIVY 374

RESULT 15

US-09-107-532A-6629  
Sequence 6629, Application US/09107532A

Patent No. 6583275  
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FABRIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:

ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02354  
COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051,571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6629:  
SEQUENCE CHARACTERISTICS:

LENGTH: 308 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (8) LOCATION 1...308  
SEQUENCE DESCRIPTION: SEQ ID NO: 6629:

US-09-107-532A-6629

Query Match 34.0%; Score 552.5; DB 4; Length 308;

Best Local Similarity 38.8%; Pred. No. 5, 6e-48;  
Matches 120; Conservative 65; Mismatches 113; Indels 11; Gaps 6;

QY 8 IARDVTEIIGKPIVATN-----KLAAGVAVAAKELMEPCSSVDRIGYSMIADAEK 63  
DB 121 FRAUGAEIVVVPVAGGSDADADIAKASAEKPEENAVILNQAPGFDPNAPETAGQ 180



Mon May 24 08:18:53 2004

us-09-931-457a-31.rat

Page 7

[illegible]

Search completed: May 18, 2004, 13:09:40  
Job time : 24 secs

Mon May 24 08:18:53 2004

US-09-931-457a-31.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 13:08:40 ; Search time 48 Seconds  
(without alignments)  
1884.062 Million cell updates/sec

Title: US-09-931-457a-31

Perfect score: 1623  
Sequence: 1 MAVERSGIAKQVTEIGTPT.....LSVLFESVREARSMTEFP 325

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1145566 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145566

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Lasting first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10G\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1623	100.0	325	9	US-09-931-457a-31
2	1622	99.9	325	12	US-10-424-599-273302
3	1488	91.7	325	12	US-10-424-599-198006
4	1453	89.5	325	9	US-09-931-457a-32
5	1389	85.6	328	12	US-10-425-114-59830
6	1381	85.1	328	12	US-10-425-114-47602
7	1380	85.0	328	12	US-10-425-114-48092
8	1375	84.7	328	12	US-10-425-114-59613
9	1364	84.0	325	12	US-10-425-114-60606
10	1364	84.0	325	12	US-10-464-811-21
11	1317	81.1	345	9	US-09-931-457a-64
12	1314	81.0	346	9	US-09-931-457a-66
13	1309	80.7	316	12	US-10-425-114-60706
14	1279	78.8	397	12	US-10-425-114-57303
15	1279	78.8	398	12	US-10-425-114-71564

16	1275	78.6	398	9	US-09-931-457a-62
17	1275	78.6	402	12	US-10-425-114-69297
18	1271	78.3	383	9	US-09-931-457a-65
19	1234	76.0	324	12	US-10-424-599-204076
20	1231	75.8	320	12	US-10-424-599-198002
21	1179	72.6	324	12	US-09-905-290A-4
22	1079	66.5	372	12	US-10-424-599-196219
23	985	60.7	212	12	US-10-425-114-45564
24	968.5	59.7	315	12	US-10-282-122A-46820
25	948.5	58.5	310	12	US-10-282-122A-62704
26	947.5	58.4	310	12	US-10-282-122A-64592
27	947.5	58.4	310	12	US-10-424-599-160879
28	941.5	58.0	310	12	US-10-282-122A-63762
29	940.5	57.9	310	12	US-10-282-122A-63092
30	917	56.5	309	12	US-10-282-122A-52203
31	916.5	56.5	309	12	US-10-282-122A-52203
32	912	56.2	310	12	US-10-282-122A-52032
33	896.5	55.2	309	12	US-10-282-122A-52032
34	882	54.3	322	12	US-10-282-122A-77194
35	879	54.2	317	12	US-10-282-122A-68735
36	870	53.6	332	12	US-10-282-122A-45004
37	865.5	53.3	323	12	US-10-282-122A-74982
38	865.5	53.3	323	12	US-10-282-122A-7586
39	865.5	53.3	323	12	US-10-464-811-22
40	863.5	53.2	323	12	US-10-282-122A-55655
41	862	53.1	322	12	US-10-282-122A-77839
42	858.5	52.9	323	9	US-09-815-242-10239
43	857.5	52.8	323	12	US-10-282-122A-56629
44	857.5	52.8	323	12	US-10-282-122A-56629
45	856	52.7	324	9	US-09-815-242-11879

#### ALIGNMENTS

RESULT 1

US-09-931-457a-31

Sequence 31, Application US/09931457A

Patent No. US20020157132A1

GENERAL INFORMATION:

APPLICANT: Falco, S. Carl

TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes

FILE REFERENCE: B1116 US CIP

CURRENT APPLICATION NUMBER: US/09/931,457A

CURRENT FILING DATE: 2002-02-22

PRIOR APPLICATION NUMBER: 09/424,976

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: 60/065,385

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/049,406

NUMBER OF SEQ ID NOS: 72

SOFTWARE: Microsoft Office 97

SEQ ID NO 31

LENGTH: 325

TYPE: PRT

ORGANISM: Glycine max

US-09-931-457a-31

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e-154;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAVERSGIAKQVTEIGTPT.....LSVLFESVREARSMTEFP 325	60
DB	1	MAVERSGIAKQVTEIGTPT.....LSVLFESVREARSMTEFP 325	60
QY	61	BEKGITGKSVLIEPTSGNTGIGLAFMAARGVLIITTPASMSLERRIILLAGAGIIV 120	120
DB	61	BEKGITGKSVLIEPTSGNTGIGLAFMAARGVLIITTPASMSLERRIILLAGAGIIV 120	120
QY	121	LTPDAKKGKAGVQKKEILLATPNAVYIQGEPNPANPVHETTCPEIWKSGDKIDAFV 180	180

```

Db      121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
Qy      241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        241 FGERYLSVLPESVRRBAESMTFEP 325
        301 FGERYLSVLPESVRRBAESMTFEP 325

```

RESULT 2

```

US-10-424-599-273302
; Sequence 273302, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273302
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88813C.1.pep
US-10-424-599-273302

```

```

Query Match          99.9%; Score 1622; DB 12; Length 325;
Best Local Similarity 99.7%; Pred. No. 1,4e-154;
Matches 324; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MAVERSGIAKDVTELIGKTPLYINKLADGCVARVAAKLMEPCSSVKDRIGYSMTADA 60
        1 MAVERSGIAKDVTELIGKTPLYINKLADGCVARVAAKLMEPCSSVKDRIGYSMTADA 60
        61 BEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
        61 BEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
        121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
        241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        301 FGERYLSVLPESVRRBAESMTFEP 325
        301 FGERYLSVLPESVRRBAESMTFEP 325

```

```

RESULT 3
US-10-424-599-198006
; Sequence 198006, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```

```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198006
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20824C.1.pep
US-10-424-599-198006

```

```

Query Match          91.7%; Score 1488; DB 12; Length 325;
Best Local Similarity 89.2%; Pred. No. 4.3e-141;
Matches 290; Conservative 21; Mismatches 14; Indels 0; Gaps 0;

```

```

Qy      1 MAVERSGIAKDVTELIGKTPLYINKLADGCVARVAAKLMEPCSSVKDRIGYSMTADA 60
        1 MAVERSGIAKDVTELIGKTPLYINKLADGCVARVAAKLMEPCSSVKDRIGYSMTADA 60
        61 BEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
        61 BEKGLITPGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRITLLAFGAEIV 120
        121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        121 LTPDAKMGKAVQKAEIETAKTLPNAYILQOFENPANKHYETTTGPEIMKSGDKIDAFV 180
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
        181 SGIGTGTTGAGKYLKEONPNKILGVEPVESPVLSGGKPGPHKIQIGAGTIPGVLEV 240
        241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        241 NLDEVOISSDEAIFETAKLALKEGLFVGISGAAAAAFOIAKREPENAKLIVAVPS 300
        301 FGERYLSVLPESVRRBAESMTFEP 325
        301 FGERYLSVLPESVRRBAESMTFEP 325

```

RESULT 4

```

US-09-931-457a-32
; Sequence 32, Application US/09931457A
; Patent No. US20020157132A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes
; FILE REFERENCE: B1116 US CIP
; CURRENT APPLICATION NUMBER: US/09/931,457A
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 09/424,976
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/065,385
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/049,406
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Citrullus lanatus
US-09-931-457a-32

```

```

Query Match          89.5%; Score 1453; DB 9; Length 325;

```

Best Local Similarity 87.1%; Pred. No. 1.5e-137;  
Matches 203; Conservative 24; Mismatches 10; Indels 0; Gaps 0;

```
QY 1 WAVERSGIAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MADASTIAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 60
QY 61 EKGITGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 EKGITGKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 120
QY 121 LTPDPAKMGKAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFV 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LTPDPAKMGKAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFV 180
QY 181 SGITGGTGTGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDV 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SGITGGTGTGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDV 240
QY 241 NLIDEVVOISDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFS 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 NLIDEVVOISDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFS 300
QY 301 FGERTYLSVLFESVRRBASMTFEP 325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 FGERTYLSVLFESVRRBASMTFEP 325
```

RESULT 5  
US-10-425-114-59830

```
; Sequence 59830, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59830
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3900-005-C7_F11.pep
US-10-425-114-59830
```

Query Match 85.6%; Score 1389; DB 12; Length 328;  
Best Local Similarity 84.9%; Pred. No. 4.2e-131;

Matches 270; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

```
QY 8 IAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 IAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 70
QY 68 PKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 127
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 PKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 130
QY 128 MKGAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFVSGIGTGS 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 MKGAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFVSGIGTGS 190
QY 188 TITGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDVNLIDEV 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 TITGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDVNLIDEV 250
```

```
QY 248 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFSFGERTYLS 307
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFSFGERTYLS 310
```

```
QY 308 SVLFESVRRBASMTFEP 325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 SVLFESVRRBASMTFEP 328
```

RESULT 6

```
US-10-425-114-47602
; Sequence 47602, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47602
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700153393_F11.pep
US-10-425-114-47602
```

Query Match 85.1%; Score 1381; DB 12; Length 328;  
Best Local Similarity 84.6%; Pred. No. 2.7e-130;

Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

```
QY 8 IAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 IAKDVTELTIGTPLYLTKADGCVARVAAKLEMEPCSSYKDRIGYMTDAEERGLIT 70
QY 68 PKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 127
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 PKSVLIEPTSGNTGIGLAFMAAARGYKLIITMPASMSLERIILAFGABLVLTDPAG 130
QY 128 MKGAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFVSGIGTGS 187
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 MKGAVQAKAEELIATKPNAYIIQOFENPANPKHYETTGDEIWKSGDKIDAFVSGIGTGS 190
QY 188 TITGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDVNLIDEV 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 TITGAGTKLEQNPNIKLYGVPESAVLSGKPPHKOIGAGFIPGVLDVNLIDEV 250
QY 248 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFSFGERTYLS 307
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGLIVAVPFSFGERTYLS 310
QY 308 SVLFESVRRBASMTFEP 325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 SVLFESVRRBASMTFEP 328
```

RESULT 7

```
US-10-425-114-49092
; Sequence 49092, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
```

APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 49092  
 LENGTH: 328  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3150-027-D8\_FLI.pep  
 US-10-425-114-49092

Query Match 85.1%; Score 1381; DB 12; Length 328;  
 Best Local Similarity 84.6%; Pred. No. 2.7e-130;  
 Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 8 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 67  
 DB 11 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 70  
 QY 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 127  
 DB 71 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 130  
 QY 128 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 187  
 DB 131 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 190  
 QY 188 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 247  
 DB 191 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 250  
 QY 248 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 307  
 DB 251 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 310  
 QY 308 SVLFESVRRERASMTPEP 325  
 DB 311 SVLFESVRRERASMTPEP 328

RESULT 8  
 US-10-425-114-59613  
 Sequence 59613, Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jinsong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 59613  
 LENGTH: 328  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3597-017-F4\_FLI.pep  
 US-10-425-114-59613  
 Query Match 85.0%; Score 1380; DB 12; Length 328;  
 Best Local Similarity 84.6%; Pred. No. 3.4e-130;  
 Matches 269; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 8 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 67  
 DB 11 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 70  
 QY 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 127  
 DB 71 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 130  
 QY 128 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 187  
 DB 131 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 190  
 QY 188 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 247  
 DB 191 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 250  
 QY 248 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 307  
 DB 251 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 310  
 QY 308 SVLFESVRRERASMTPEP 325  
 DB 311 SVLFESVRRERASMTPEP 328

RESULT 9  
 US-10-425-114-60606  
 Sequence 60606, Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jinsong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 60606  
 LENGTH: 328  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3597-067-G10\_FLI.pep  
 US-10-425-114-60606

Query Match 84.7%; Score 1375; DB 12; Length 328;  
 Best Local Similarity 84.3%; Pred. No. 1.1e-129;  
 Matches 268; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

QY 8 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 67  
 DB 11 IADVTLLIGKTPVLYLNKLDGCVARVAANKLMEPCSSVNDRIIGYSMTADAEKGLIT 70  
 QY 68 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 127  
 DB 71 PKGSVLIPTSGNTGIGLAFAAARGYKLIITMPASMSERRITLLAFGAELVLPDPAK 130  
 QY 128 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 187  
 DB 131 MKGAVOKAEELIATKPNAYIILOFENPANKVHYETTPGRIWKSDDKIDAFVSGIGTGG 190  
 QY 188 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 247  
 DB 191 TTGAGKTYLKEQNPNIKIGVEPVESPVLSGGKPGPHKIQIGAGFIPGLVENVLLDEVY 250  
 QY 248 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 307  
 DB 251 QISSDEALITAKLALKEGLFVGISGGAATAAFAQIAKRPENAGKLIIVAVPSPGERRYLS 310

QY 308 SVLFESVREAREASMTPEP 325  
 DB 311 SVLFOSIKKEAESMVEP 328

RESULT 10  
 US-10-464-811-21

Sequence 21, Application US/10464811  
 Publication No. US20040033219A1  
 GENERAL INFORMATION:  
 APPLICANT: Kraus, Jan  
 APPLICANT: Oliveriusova, Jana  
 TITLE OF INVENTION: Human Cystathionine B-Synthase Variants and Methods of Production  
 FILE REFERENCE: 2848-49  
 CURRENT APPLICATION NUMBER: US/10/464,811  
 PRIOR FILING DATE: 2003-06-17  
 PRIOR APPLICATION NUMBER: 60/389,541  
 PRIOR FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Patent version 3.1  
 SEQ ID NO 21  
 LENGTH: 325  
 TYPE: PRT  
 ORGANISM: Triticum aestivum  
 US-10-464-811-21

Query Match 84.0%; Score 1364; DB 12; Length 325;  
 Best Local Similarity 82.3%; Pred. No. 1,4e-128;  
 Matches 261; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 8 IAKDVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 67  
 DB 9 IAKDVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 68  
 QY 68 PKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLTDPAG 127  
 DB 69 PKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLTDPAG 128  
 QY 128 MKGAVOKAEELIAKTPENAYILQGFENPANKVHYETTGEIWKSGDKIDAFVSGIGTG 187  
 DB 129 MKGAVOKAEELIAKTPENAYILQGFENPANKVHYETTGEIWKSGDKIDAFVSGIGTG 188  
 QY 188 TTTGAKRYLKEONPNKILGVEPVESPVLSGKRGPHKIOGIGAGFIPGVLEVNILDEV 247  
 DB 189 TTTGAKRYLKEONPNKILGVEPVESPVLSGKRGPHKIOGIGAGFIPGVLEVNILDEV 248  
 QY 248 QISSDEALETAKLALKEGLFVIGISSGAAAAAFQIAKPEPNAKGLIVAVPSPFGRYLS 307  
 DB 249 QISSDEALETAKLALKEGLFVIGISSGAAAAAFQIAKPEPNAKGLIVAVPSPFGRYLS 308  
 QY 308 SVLFESVREAREASMTPEP 324  
 DB 309 SVLFOSIKKEAESMVEP 325

RESULT 11  
 US-09-931-457a-64

Sequence 64, Application US/09931457A  
 Patent No. US20020157132A1  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Allen, Stephen M.  
 TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
 FILE REFERENCE: B0116 US CIP  
 CURRENT APPLICATION NUMBER: US/09/931,457A  
 PRIOR FILING DATE: 2002-02-22  
 PRIOR APPLICATION NUMBER: 09/424,976  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: 60/065,385  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/049,406  
 PRIOR FILING DATE: 1997-06-12

NUMBER OF SEQ ID NOS: 72  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 64  
 LENGTH: 415  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 US-09-931-457a-64

Query Match 81.1%; Score 1317; DB 9; Length 415;  
 Best Local Similarity 77.4%; Pred. No. 1.1e-123;  
 Matches 250; Conservative 39; Mismatches 34; Indels 0; Gaps 0;

QY 3 VERSGIADVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 62  
 DB 93 VERSGIADVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 152  
 QY 63 KGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLT 122  
 DB 153 KGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLT 212  
 QY 123 DPAKMGAVOKAEELIAKTPENAYILQGFENPANKVHYETTGEIWKSGDKIDAFVSG 182  
 DB 213 DPAKMGAVOKAEELIAKTPENAYILQGFENPANKVHYETTGEIWKSGDKIDAFVSG 272  
 QY 183 IGTGTTGAKRYLKEONPNKILGVEPVESPVLSGKRGPHKIOGIGAGFIPGVLEVNILDEV 242  
 DB 273 IGTGTTGAKRYLKEONPNKILGVEPVESPVLSGKRGPHKIOGIGAGFIPGVLEVNILDEV 332  
 QY 243 LDEVVQISSDEALETAKLALKEGLFVIGISSGAAAAAFQIAKPEPNAKGLIVAVPSPFG 302  
 DB 333 LDEVVQISSDEALETAKLALKEGLFVIGISSGAAAAAFQIAKPEPNAKGLIVAVPSPFG 392  
 QY 303 ERYLSSVLFESVREAREASMTPEP 325  
 DB 393 ERYLSSVLFESVREAREASMTPEP 415

RESULT 12  
 US-09-931-457a-66

Sequence 66, Application US/09931457A  
 Patent No. US20020157132A1  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Allen, Stephen M.  
 TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes  
 FILE REFERENCE: B0116 US CIP  
 CURRENT APPLICATION NUMBER: US/09/931,457A  
 PRIOR FILING DATE: 2002-02-22  
 PRIOR APPLICATION NUMBER: 09/424,976  
 PRIOR FILING DATE: 1999-12-02  
 PRIOR APPLICATION NUMBER: 60/065,385  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/049,406  
 PRIOR FILING DATE: 1997-06-12  
 NUMBER OF SEQ ID NOS: 72  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 66  
 LENGTH: 386  
 TYPE: PRT  
 ORGANISM: Solanum tuberosum  
 US-09-931-457a-66

Query Match 81.0%; Score 1314; DB 9; Length 386;  
 Best Local Similarity 78.1%; Pred. No. 1.9e-123;  
 Matches 249; Conservative 38; Mismatches 32; Indels 0; Gaps 0;

QY 3 VERSGIADVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 62  
 DB 64 VERSGIADVTEHIGKTPVLYNKLADGCVARVAAKELMEPPCQSVKORIGYSMTADAE 123  
 QY 63 KGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLT 122  
 DB 124 KGLITPGKSVLIEPTSGNTGIGLAFMAARGYKLIITPASMSEERRIILAFGAEVLT 183

Mon May 24 08:18:53 2004

us-09-931-457a-31.rapb

Page 6

123 DPAKMGKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 182  
184 DPAKMGKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 243  
183 IGTGTTTGTAGKYLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 242  
244 IGTGTTTGTAGKYLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 303  
243 LDEVOISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 302  
304 MDEYIEISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 363  
303 ERYLSSVLFESVRRAESM 321  
364 ERYLSSVLFESVRRAESM 382

RESULT 13

US-10-425-114-60706  
Sequence 60706, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack B.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 60706  
LENGTH: 316  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3587-229-H7\_FLI.pep  
US-10-425-114-60706

Query Match 80.7%; Score 1309; DB 12; Length 316;  
Best Local Similarity 81.1%; Pred. No. 4,56-123;  
Matches 258; Conservative 22; Mismatches 26; Indels 12; Gaps 1;

8 IAKDTEILIGTPELVYLNKADCCVARYAAKLEIMPCSSVYKDRIGYIMIDAEKGLIT 67  
11 IAKDTEILIGTPELVYLNKADCCVARYAAKLEIMPCSSVYKDRIGYIMIDAEKGLIT 58  
68 PCKSVLIEPTSGNTGIGAFAPAAAGYKLIITMPASMSLEIRIILAFGALVLTDPAG 127  
59 PCKSVLIEPTSGNTGIGAFAPAAAGYKLIITMPASMSLEIRIILAFGALVLTDPAG 118  
128 MKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 187  
119 MKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 178  
188 TTTGAGTKLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 247  
179 TTTGAGTKLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 238  
248 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 307  
229 QISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 298  
308 SVLFESVRRAESM 325  
299 SVLFESVRRAESM 316

RESULT 14

US-10-425-114-57303  
Sequence 57303, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack B.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 57303  
LENGTH: 397  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLMO17258D02\_FLI.pep  
US-10-425-114-57303

Query Match 78.8%; Score 1279; DB 12; Length 397;  
Best Local Similarity 74.9%; Pred. No. 6,7e-120;  
Matches 245; Conservative 44; Mismatches 36; Indels 2; Gaps 1;

1 MAVERSG-IKADTEILIGTPELVYLNKADCCVARYAAKLEIMPCSSVYKDRIGYIMIDAEKGLIT 58  
71 VAAEVLGGLNTANMTOILGNTPMVYLNKAVGSAVAAALKEIMPCSSVYKDRIGYIMIDAEKGLIT 130  
59 DAEKGLITGKSVLIEPTSGNTGIGAFAPAAAGYKLIITMPASMSLEIRIILAFGALVLTDPAG 118  
131 DAEKGLITGKSVLIEPTSGNTGIGAFAPAAAGYKLIITMPASMSLEIRIILAFGALVLTDPAG 190  
119 LVLTDPAGKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 178  
191 LVLTDPAGKGAQKAEIILAKTNPAYILQFPENPANKVHYETTGTGEIWKSGDKIDAVSG 250  
179 FVSGIGGTTTGTAGKYLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 238  
251 FVSGIGGTTTGTAGKYLKEQNPNTKLGVEPVSPTSGKRGPHKIOGIGAGTIPVLEVL 310  
239 EYVILDEVOISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 298  
311 EYVILDEVOISSDEALETAKLALKEGLFVGISGAAAAAFOIARPENAGKLIYAVFPSPG 370  
299 PSFGERYLSVLFESVRRAESM 325  
371 PSFGERYLSVLFESVRRAESM 397

RESULT 15

US-10-425-114-71564  
Sequence 71564, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jindong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack B.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 71564  
LENGTH: 398  
TYPE: PRT

Mon May 24 08:18:53 2004

us-09-931-457a-31.rapb

Page 7

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLM017281G01\_FLI.pep  
US-10-425-114-71564

Query Match 78.8%; Score 1279; DB 12; Length 398;  
Best Local Similarity 74.9%; Pred. No. 6,7e-120;  
Matches 245; Conservative 44; Mismatches 36; Indels 2; Gaps 1;

```
QY 1 MAVRRSG--IAKQVTEHIGKTEPLVYLNKLDGCVARVAARKLEMEPCSSVQDRIGYSMTA 58
Db 72 VAAVGGALNANDVTQLIGNTPMYLNNAVKAQSVANVAARKLEMEPCSSVQDRIGYSMTN 131
QY 59 DAEKGLITTPGKSVLIPTSGNTGIGLAFMAAARGYKLIITMPASMSLERRIILAFGAE 118
Db 132 DAEKGLITTPGKSVLVEATSGNTGIGLAFMAASKGYKLIITMPSSMSMERVILAFGAE 191
QY 119 LVLTDPKAKMGAVQKAEHILAKTPNAYILQGFNPAMPKHYETTGPEITMKGSDKIDA 178
Db 192 LVLTDAKMGKALDKATEILNKTPNSTYMLQGFNPAMPKHYETTGPEITMKGSKYDI 251
QY 179 FVSGIGTGTITGAGKYLKQNPNIKLIGVFPVSPVLSGKPGPHKIQIGAGFIPGV 238
Db 252 FVSGIGTGTITGAGKYLKQNPNIKLIGVFPVSPVLSGKPGPHKIQIGAGFIPRV 311
QY 239 EVNLDEVOISSDEALETALALKEGLFVGISSGAAAAAFOIAKRPENAGKLIYVVF 298
Db 312 DSDILDEVIEISSDEAVETAKOLAVQEGILVGISSGAAAAAIVAKRPENAGKLIYVVF 371
QY 299 PSFGRYLSVLPFSVREAREASMTPEP 325
Db 372 PSFGRYLSVLPFSVREAREASMTPEP 398
```

Search completed: May 18, 2004, 13:14:30  
Job time : 49 secs